ON CONCEPT CLASSES WITH LOW NO-CLASH TEACHING DIMENSION

A Thesis

Submitted to the Faculty of Graduate Studies and Research

In Partial Fulfillment of the Requirements

for the Degree of

Master of Science

In

Computer Science

University of Regina

By

Abolghasem Soltani

Regina, Saskatchewan

May 2021

UNIVERSITY OF REGINA

FACULTY OF GRADUATE STUDIES AND RESEARCH

SUPERVISORY AND EXAMINING COMMITTEE

Abolghasem Soltani, candidate for the degree of Master of Science in Computer Science**,** has presented a thesis titled, ***On Concept Classes with Low No-Clash Teaching Dimension,*** in an oral examination held on April 30, 2021. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner: *Dr. Nader Mobed, Department of Physics

Co-Supervisor: *Dr. Sandra Zilles, Department of Computer Science

Co-Supervisor: *Dr. Shaun Fallat, Department of Mathematics & Statistics

Committee Member: *Dr. Boting Yang, Department of Computer Science

Chair of Defense: *Dr. Ian Coulson, Department of Geology

*via ZOOM Conferencing

Abstract

Computational Learning Theory studies the complexity of learning for various formal models of machine learning. Such models use a learning algorithm $\mathcal{A}$ fed with a set of labelled data for a target concept belonging to a given class of concepts, and $\mathcal{A}$ produces a mapping that is used to accurately predict the correct labels of unseen instances. Machine teaching is concerned with settings in which a teacher provides the learners with helpfully chosen data.

Different learning constraints, such as the way a teacher and a learner interact with each other, produce a variety of teaching models for which one task is to answer the question "what is the minimum number of labelled examples required by the learning algorithm $\mathcal{A}$ to exactly identify any given concept C belonging to $\mathcal{C}$?".

In this thesis, we use combinatorics to study two well-known notions of teaching that avoid unwanted forms of collusion between the teacher and the learner, namely the *recursive teaching dimension* (RTD) and the *no-clash teaching dimension* (NCTD). We affirmatively answer the question whether there exists a concept class

for which $\frac{\text{RTD}}{\text{NCTD}} > 2$ holds. Along these lines, the smallest concept class for which the equality RTD $= 3 \cdot$ NCTD holds is provided. Using design theory, we construct a concept class of size 16 over 8 instances that satisfies RTD $= 4$ and NCTD $= 1$. This is the largest multiplicative gap found between RTD and NCTD so far. Furthermore, by implementing a Python toolbox to compute and compare teaching complexity parameters, we succeed to generate numerous concept classes over at most five instances for which our well-known ratio is at least 3.

Over 4 instances, it is shown that the smallest concept class which satisfies RTD $=$ $3 \cdot$ NCTD is unique. Over 5 instances, there are three non-equivalent concept classes of size 8, and also of size 9 for which RTD $= 3$ and NCTD $= 1$ hold. Additional concept classes of sizes 10, 12, 14, and 16 are also generated with the property RTD $=$ $3 \cdot$ NCTD.

Finally, by encoding certain teaching information in the one-inclusion graph, we establish a connection between the one-inclusion graph of concept classes and the non-clash teaching complexity. We find that for the so-called shortest-path-closed concept classes with NCTD equal to 1, the corresponding one-inclusion graphs have at most one cycle, and consequently the recursive teaching dimension is at most 2.

Acknowledgments

I am deeply indebted to my supervisors Dr. Sandra Zilles and Dr. Shaun Fallat for their constant support, encouragement and consideration throughout my MSc program. The opportunity that you gave me was a turning point in my life and I will never forget your great attitude and kindness.

I would like to acknowledge Dr. David Kirkpatrick for his collaboration through my research work. I learned from David that sometimes one single word may change everything. His magic word was *Geometry*. I also appreciate the recommendations and hints offered by Dr. Karen Meagher during my research.

I would like to express my sincere appreciation to Dr. Boting Yang and Dr. Nader Mobed for examining my thesis.

I also gratefully acknowledge the financial support from my supervisors which made this work possible. I really thank the Department of Computer Science and the Department of Mathematics and Statistics for given me an opportunity to work as a teaching assistant.

Dedication

It is my genuine gratefulness and warmest regard that I dedicate this work to my wife *Ziba* and my daughter *Emma* who have brought so much love, joy, and happiness into my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Computer systems need a mechanism for learning once a proposed problem is not being solved directly by writing an algorithm but by employing example data or past experience. Learning is required when there is a lack of knowledge or the problem changes over time, such as in recognition of spoken speech or in sales prediction, respectively. Such problems are investigated in machine learning. *Machine Learning* refers to the study of systems in which a predictive model is obtained from data [1]. *Machine Teaching* is a subfield concerned with the design of an optimal training set or a teaching set for the learner. In other words, it studies the interaction between a teacher and a learner in which the teacher can provide information about a target concept by selecting labeled examples. A concept is simply a subset of a given instance space, and a labeled example (x, l) for a concept C consists of an instance x in the instance space and a label $l \in \{0, 1\}$ that indicates whether or not x belongs to C.

1.1 Notions of Teaching Dimension

One important research direction in machine teaching is *Algorithmic Teaching Theory*, which is a branch of *Computational Learning Theory* and studies a well-known notion in learning theory called *Teaching Dimension*. This quantity is defined to be the worst-case number of examples needed to be selected by the teacher in the process of teaching any of the potential target concepts [28, 39]. Formally, in a teaching model, the teacher is a mapping that produces a finite set of labelled examples for each concept in a given concept class, i.e., set of potential target concepts. The details, such as which labelled examples are given to the learner or when learning is considered successful, all depend on the selected model. Such models commonly include a set of rules which governs the exchange of information between the teacher and the learner. Different notions of teaching dimension are provided by different teaching models.

In the oldest teaching model, called the *classical teaching model*, introduced by Goldman and Kearns [16], both teacher and learner know the concept class, and the teacher produces a teaching set of minimum size that distinguishes a concept from all other concepts in the class. The teaching complexity (called teaching dimension) refers to the teaching set size for the most difficult to teach concept in the class.

Changing the machine teaching notion leads to new teaching models whose teaching complexity will be at most the classical teaching dimension (TD). Zilles et al. [40] and Gao et al. [15] have proposed the notions of *Recursive Teaching Dimension* (RTD) and *Preference-based Teaching Dimension* (PBTD), respectively, in which the learner is aware that it is being taught by a teacher. While both proposed teaching models employ fewer examples for teaching than the classical teaching model, over a concept class of infinite size, the preference-based teaching model outperforms the recursive teaching model. For more details on this model, the reader is referred to [15].

Generally, a teaching model must impose constraints in order to avoid "unfair coding tricks" or "collusion" between the teacher and the learner while they share information. The most widely accepted notion of collusion was provided by Goldman and Mathias in 1996 [17] and has led to the design of various teaching models satisfying a corresponding condition of collusion-freeness. The condition stipulates, for any concept class $\mathcal{C}$ and for any concept $C \in \mathcal{C}$: (i) the examples in the teaching set $T(C)$ for C are labelled consistently with C, and (ii) on any super-set S of $T(C)$ which is consistent with C, the learner will output C.

Recently Kirkpatrick et al. [24] proposed a new teaching model called *no-clash teaching*, together with the corresponding teaching dimension parameter NCTD, which is provably the smallest teaching complexity that any collusion-free teaching models can have.

In all the teaching models mentioned above, the learner is provided with a training set by a teacher and then the learner outputs a concept. The order of training examples in the teaching set plays no role. Such models are called *Batch Teaching*

Models. By contrast, motivated by the fact that humans respond better to sequential information, a teacher may optimize the order of training items in order to have a more efficient interaction with the learner. This results in other teaching models, named *Sequential Teaching Models*, which are useful in teaching robots [28, 39].

In direct contrast to teaching models in which learners are given helpfully chosen examples, another learning model to point out here is a model called the *Probably Approximately Correct* (PAC) learning model, proposed by Valiant in 1984 [38], in which the learner is given examples randomly drawn from an underlying distribution. In PAC-learning, sample complexity is characterized by a combinatorial parameter called the *Vapnik-Chervonenkis Dimension* (VCD) of the concept class [2, 37].

1.2 Motivation

Characterizing sample complexity by the VCD of a given concept class raises the question whether there is any type of relationship between the well-studied teaching dimension parameters and the VCD for a given concept class.

There is no general relationship between TD and VCD in all teaching settings. However, for any finite concept class $\mathcal{C}$, so far, some notable results have been reported in the literature regarding RTD. Researchers have been particularly interested in the question whether the RTD is upper-bounded by a function that grows only linearly in the VCD [31].

Doliwa et al. in [7] investigated concept classes which fulfill $\text{RTD}(\mathcal{C}) \leq \text{VCD}(\mathcal{C})$ and concept classes which satisfy $\text{RTD}(\mathcal{C}) > \text{VCD}(\mathcal{C})$. Moreover, the best known worst-case lower bound is $\text{RTD}(\mathcal{C}) \geq \frac{5}{3}\text{VCD}(\mathcal{C})$, which is proposed and proven by Chen et al. in 2016 [4]. On the other hand, to date the best known upper bound is due to Hu et al. (2017) [21], who proved that for a given concept class $\mathcal{C}$, $\text{RTD}(\mathcal{C}) \leq 39.3752\ \text{VCD}(\mathcal{C})^2 - 3.6330\ \text{VCD}(\mathcal{C})$, and consequently, since NCTD is the optimal teaching complexity, it follows that NCTD is at most quadratic in VCD.

In contrast to the best known complexity result for batch models (NCTD) which is quadratic in the VC dimension, Mansouri et al. in their recent work [28] identified a novel sequential model with teaching complexity linear in the VC dimension of the concept class.

Bounds on teaching parameters in terms of the VCD are worthwhile to consider because they suggest a close relationship between learning from randomly chosen examples and learning from teachers, in terms of information complexity [7, 8]. It may bring new insights into the "sample compression conjecture" which claims that each sample set S consistent with a concept in $\mathcal{C}$ can be compressed to a subset of size at most the VCD of $\mathcal{C}$ without losing any label information [12, 27].

These research works motivate us to continue studying in this same area of research and investigate those hypothesis classes that reveal more information about the strongest collusion-free batch model, i.e. the non-clashing model.

1.3 Major Contributions and Organization

While it has been theoretically shown by Kirkpatrick et al. [24] that the no-clash teaching model is optimal among all models satisfying Goldman and Mathias's notion of collusion-freeness, so far we only were aware of concept classes for which the ratio $\frac{\text{RTD}}{\text{NCTD}}$ is at most 2. This raises the question about how large can the ratio between RTD and NCTD be - the main question addressed in this thesis.

Once some essential definitions regarding concept classes are provided and teaching parameters from the literature have been introduced in Chapter 2, the main focus of Chapter 3 is to address the question whether there exists a concept class for which $\frac{\text{RTD}}{\text{NCTD}} > 2$ holds. By providing a few new examples of concept classes, we will show that the answer to this question is positive. Along with this matter, we have made use of the wealth of combinatorial structure from design theory and linked it with algorithmic teaching theory. During the past decades, design theory has found numerous applications and many researchers have been interested in working on specific combinatorial designs, such as balanced incomplete block designs, Hadamard matrices, orthogonal arrays, etc., which has led to powerful new combinatorial and computational techniques [33]. This work may take a first step in creating synergies between learning theory and design theory in order to finding new results regarding teaching models in terms of sample complexity.

The last section of Chapter 3 describes the implementation of a python toolbox to compute teaching complexity parameters. The toolbox helps us to compute such parameters for a large number of concept classes. In addition, analyzing those concept classes whose structure is more sophisticated is eased by this toolbox. We found that over 4 instances, there is a unique concept class of size 8 which fulfills RTD $=3$ and NCTD $=1$. On the other hand, over 5 instances, there are three non-equivalent concept classes of size 8, and also of size 9, as well as one unique concept class of size 10 for which RTD $=3$ and NCTD $=1$ hold. We generated concept classes of size 12, 14 and 16 over domains of size 6, 7 and 8, respectively, for which RTD $=3\cdot$ NCTD. Moreover, we found a concept class of size 16 over 8 instances that satisfies RTD $=4$ and NCTD $=1$, and this is the biggest multiplicative gap observed between RTD and NCTD so far.

To continue, Chapter 4 opens a discussion about graph-theoretic representations of concept classes in order to extract information about the teachability of concepts in concept classes. Although the learner can distinguish a target concept from other concepts using the associated teaching sets, the useful distinguishing information encoded in teaching sets is not always easy to observe by the learner in the matrix representation of a concept class. The *one-inclusion graph* of a concept class, introduced by Bondy in 1972 [3], is a helpful graph-theoretic representation for studying some notions in learning theory [11, 18–20]. Research works such as [9, 25] have witnessed some links between the one-inclusion graphs of certain concept classes and their corresponding teaching complexity including VC-dimension, classical and recursive teaching dimensions.

Motivated by such investigations above, the main result of Chapter 4 features a connection between the one-inclusion graph of concept classes and the non-clash teaching complexity. We prove that for so-called shortest-path-closed concept classes with NCTD equal to 1, the associated one-inclusion graphs have at most one cycle, and consequently the recursive teaching dimension is at most 2.

Finally, we conclude our work in Chapter 5, providing open problems and conjectures related to the various concepts discussed in this thesis.

Chapter 2

Preliminaries

This chapter reviews three teaching models employed in our work: the *teaching dimension* model, the *recursive teaching* model, and the recently defined *no-clash teaching* model. In addition, out of numerous facts discovered by researchers in this area, some examples and theorems will be highlighted in order to pave the way for explaining our new results. At the end of the chapter, we provide some basic definitions typically used in graph theory.

2.1 Teaching Models

In a formal learning process, consider a *domain set* $\mathcal{X}$ of size m containing instances $x_1, \ldots, x_m$. A *concept class* $\mathcal{C}$ over $\mathcal{X}$ is defined to be a subset of $\{C : C \subseteq \mathcal{X}\}$. Each element of this set is called a *concept* over $\mathcal{X}$. Equivalently, the binary representation allows us to identify a concept C over $\mathcal{X}$ by mapping x to 1 if and only if $x \in C$.

Therefore, a concept is an ordered list $C : (l_1, \ldots, l_m)$ of size m such that $l_i \in \{0, 1\}$, $1 \leq i \leq m$. By flipping the labels l_i in the ordered list C from 0 to 1 and vice versa, the complement of C, denoted by $\hat{C}$, is formed. As an example, $(1, 0, 0, 1)$ over the domain $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ features the concept $C = \{x_1, x_4\}$, and $\hat{C} = \{x_2, x_3\}$ corresponds to the ordered list $(0, 1, 1, 0)$.

A *sample set* S, containing labelled examples (x, l) where $x \in \mathcal{X}$ and $l \in \{0, 1\}$, is consistent with a concept C if for every pair $(x, l) \in S$ we have that $C(x)$ is equal to the label l of x. Moreover, we set $\mathcal{X}(S) = \{x : (x, 0) \in S \text{ or } (x, 1) \in S\}$. The sample set S is called a *teaching set* for C (with respect to $\mathcal{C}$) if C is the only concept in the concept class $\mathcal{C}$ consistent with S. In addition, A *teacher mapping* for $\mathcal{X}$ is a mapping T on $\mathcal{C}$ such that for all $C \in \mathcal{C}$, $T(C)$ is a finite sample set consistent with C.

Remark 2.1.1 *For convenience, when $\mathcal{X}$ is finite, it is often helpful to use a matrix structure to represent a concept class $\mathcal{C}$, formalized as follows. Consider a concept class $\mathcal{C} = \{C_1, \ldots, C_n\}$ over a domain $\mathcal{X} = \{x_1, \ldots, x_m\}$, m, $n \in \mathbb{N}$. Then the n by m matrix M of $\mathcal{C}$ is defined as $[C_1(\mathcal{X}) \;\; C_2(\mathcal{X}) \;\; \cdots \;\; C_n(\mathcal{X})]^T$, where the i^{th} row and the j^{th} column of M are labelled with $C_i \in \mathcal{C}$ and $x_j \in \mathcal{X}$, respectively, and the (i, j)-entry is 1 if $x_j \in C_i$ and 0 otherwise.*

The first parameter here to review is the teaching dimension of a concept class associated with the classical teaching model cited from [16, 32]:

Definition 2.1.2 *The size of the smallest teaching set for C with respect to $\mathcal{C}$ is referred to as the teaching dimension of C in $\mathcal{C}$, and is denoted by* $\mathrm{TD}(C, \mathcal{C})$. *The teaching dimension of $\mathcal{C}$ is then defined as* $\mathrm{TD}(\mathcal{C}) = \sup\{\mathrm{TD}(C, \mathcal{C}) \mid C \in \mathcal{C}\}$. *In addition,* $\mathrm{TD}_{\min}(\mathcal{C}) = \min\{\mathrm{TD}(C, \mathcal{C}) \mid C \in \mathcal{C}\}$.

The positive teaching dimension of $\mathcal{C}$, denoted by $\mathrm{TD}^+(\mathcal{C})$, is defined analogously to $\mathrm{TD}(\mathcal{C})$, where the teaching sets in the definition above contain only positive examples.

It can be seen that for distinguishing all concepts in the classical model, we need to return to the most difficult concept to teach. Therefore, it is reasonable to start the learning process with the easiest concept of the class to improve the teaching complexity obtained from TD model. Thus, the second teaching model, called the *recursive teaching model*, is introduced which is a complexity parameter obtained from cooperating between teachers and learners. By starting from a concept with the minimum teaching dimension and then removing it from the concept class, a concept with the minimum teaching dimension in the remaining class is chosen to teach and remove, and the process continues until no concepts remain in the class. Finally, the recursive teaching dimension, denoted by RTD, of the class is the largest minimum teaching dimension occurring in the process. The theoretical definition can be found in [40]:

Definition 2.1.3 *Let $\mathcal{C}$ be a concept class. Set $\mathcal{C}_0 = \mathcal{C}$, $\hat{\mathcal{C}}_0 = \emptyset$, for all $i \in \mathbb{N}$ define*

$$\hat{\mathcal{C}}_{i+1} = \{C \in \mathcal{C}_i \mid \mathrm{TD}(C, \mathcal{C}_i) = \mathrm{TD}_{\min}(\mathcal{C}_i)\},$$

and also $\mathcal{C}_{i+1} = \mathcal{C}_i \setminus \hat{\mathcal{C}}_{i+1}$. Then $\mathrm{RTD}(\mathcal{C}) = \max_{i \in \mathbb{N}} \mathrm{TD}_{\min}(\mathcal{C}_i)$ is the recursive teaching dimension of $\mathcal{C}$. By replacing TD and $\mathrm{TD}_{\min}(\mathcal{C}_i)$ with TD^+ and $\min_{C \in \mathcal{C}_i} \mathrm{TD}^+(C, \mathcal{C}_i)$, respectively, the parameter $\mathrm{RTD}^+(\mathcal{C})$ is defined similarly.

There is another well-known teaching model called *preference-based teaching model*, introduced by Gao et al. in [15], which behaves in a similar way to the recursive teaching model over any finite domain $\mathcal{X}$. However, once the domain is infinite, the preference-based teaching complexity is often smaller than the recursive teaching complexity. The preference-based teaching dimension of the concept class $\mathcal{C}$ is denoted by $\mathrm{PBTD}(\mathcal{C})$, and similarly, the positive dimension is denoted by $\mathrm{PBTD}^+(\mathcal{C})$. Throughout this thesis, all concept classes are defined on a finite domain $\mathcal{X}$, and therefore $\mathrm{PBTD}/\mathrm{PBTD}^+$ coincides with $\mathrm{RTD}/\mathrm{RTD}^+$.

The following simple example illustrates these teaching complexities.

Example 2.1.4 *Let $\mathcal{C} = \{\{x_i\} \mid i = 1, \cdots, 4\} \cup \{\phi\}$ be a concept class over a domain $\mathcal{X} = \{x_1, \cdots, x_4\}$, represented in Table 2.1. Then $\{(x_i, 1)\}$ is a teaching set for $\{x_i\}$ and therefore $\mathrm{TD}(\{x_i\}, \mathcal{C}) = 1$, $1 \leq i \leq 4$. On the other hand, to distinguish the empty concept we need to consider all 4 zero-labelled examples and so*

	x_1	x_2	x_3	x_4	TD	RTD
C_1	**1**	0	0	0	1	1
C_2	0	**1**	0	0	1	1
C_3	0	0	**1**	0	1	1
C_4	0	0	0	**1**	1	1
C_5	0	0	0	0	4	0

Table 2.1: TD = 4 and RTD = 1.

	x_1	x_2	x_3	TD	RTD
C_1	1	0	**1**	1	1
C_2	**1**	1	0	2	1
C_3	0	**1**	0	2	1
C_4	0	0	0	2	1

Table 2.2: TD = 2 and RTD = 1.

$\mathrm{TD}(\{\phi\}, \mathcal{C}) = 4$. *As a result* $\mathrm{TD}(\mathcal{C}) = 4$. *In the recursive teaching model, the easiest concepts are the first four concepts. So after removing them we just have the empty concept that does not need any teaching set. This means its teaching set cardinality is zero. Thus* $\mathrm{RTD}(\mathcal{C}) = 1$.

Regarding Table 2.2, by following the highlighted labels and the same discussion as above, the desired result for the new mentioned concept class over domain $\{x_1, x_2, x_3\}$ *are obtained.*

The other complexity parameter, which is of interest because of its connection to the above mentioned teaching complexities is the *Vapnik-Chervonenkis dimension* (VCD) of a concept class which is well-known in the PAC-learning model [37]. Consider any subset $\mathcal{X}'$ of $\mathcal{X}$ and let $C|_{\mathcal{X}'}$ be the projection of concept C on $\mathcal{X}'$. $\mathcal{X}'$ is said to be shattered by $\mathcal{C}$ if $|\{C|_{\mathcal{X}'} : C \in \mathcal{C}\}| = 2^{|\mathcal{X}'|}$. For a concept class $\mathcal{C}$, $\mathrm{VCD}(\mathcal{C})$ is the maximum size of a shattered subset of $\mathcal{X}$. In Table 2.1 let $\mathcal{X}' = \{x_1\}$ and so $\mathrm{VCD}(\mathcal{C}) = 1$, and for Table 2.2 consider $\mathcal{X}' = \{x_1, x_2\}$, which implies $\mathrm{VCD}(\mathcal{C}) = 2$.

In order to answer the question whether we have an optimal teaching model, authors in [24] have recently defined and investigated a new model of teaching called the *no-clash teaching model*, which utilizes the smallest worst-case number of examples for distinguishing concepts from each other in the underlying concept class $\mathcal{C}$. Its teaching complexity is denoted by NCTD and called the *no-clash teaching dimension*. The positive teaching complexity of this model, which works based on the positive label instances, has been introduced similarly and is denoted by NCTD^+. This new model has brought forth new research questions on comparing various complexity parameters, such as RTD and VCD, of a given concept class.

The following are some required definitions.

Definition 2.1.5 *Let T be a teacher mapping on a concept class $\mathcal{C}$. T is a non-clashing teacher if and only if there are no two distinct $C, C' \in \mathcal{C}$ such that both $T(C)$ is consistent with C' and $T(C')$ is consistent with C.*

Theorem 8 in [24] shows that a teacher mapping T is non-clashing over $\mathcal{C}$ if and only if a learner mapping L exists such that the pair (T, L) is successful and collusion-free. Therefore, the no-clash teaching dimension is defined as follows:

Definition 2.1.6 *Let $\mathcal{C}$ be a concept class over the domain $\mathcal{X}$ and T a non-clashing teacher mapping. The no-clash teaching dimension of $\mathcal{C}$ is defined and denoted as*

$$\text{NCTD}(\mathcal{C}) = \min\{\text{ord}(T, \mathcal{C}) : T \textit{ is a non-clashing teacher mapping for } \mathcal{C}\},$$

where $\text{ord}(T, \mathcal{C})$ *is the order of T on $\mathcal{C}$ and is equal to* $\sup\{|T(C)| : C \in \mathcal{C}\}$. $\text{NCTD}^+(\mathcal{C})$ *is similarly defined.*

	x_1	x_2	x_3	x_4	x_5	x_6
C_1	(**1**)	1	**1**	0	0	0
C_2	0	(**1**)	1	**1**	0	0
C_3	0	0	(1)	1	**1**	0
C_4	0	0	0	(1)	1	**1**
C_5	**1**	0	0	0	(1)	1
C_6	1	**1**	0	0	0	(1)

Table 2.3: A concept class with RTD = 2, but NCTD = $\text{NCTD}^+ = 1$.

Example 2.1.7 *The concept class in Table 2.3 has (positive) non-clash teaching dimension of size 1; the teaching sets of its concepts are circled. It is easy to see that no concept in this class has a teaching dimension of* 1 *which means* RTD ≥ 2*. On the other hand, considering the blue labels it follows that* RTD ≤ 2*.*

Remark 2.1.8 *By flipping every 0 to a 1 and vice versa in one or a few columns of the concept class* $\mathcal{C}$*, the values of* TD($\mathcal{C}$), RTD($\mathcal{C}$) *and* NCTD($\mathcal{C}$) *will not change, in particular, the sizes of all concepts teaching sets are fixed. However, the values of the corresponding positive parameters may be modified. In Table 2.3, the sample set* $S = \{(x_1, 1)\}$ *is a non-clash teaching set for the concept* C_1*. Flipping zeros and ones in the column* $[1\ 0\ 0\ 0\ 0\ 1\ 1]^T$ *associated to the instance* x_1*, results in the column* $[0\ 1\ 1\ 1\ 1\ 0\ 0]^T$ *and the sample set* $S' = \{(x_1, 0)\}$ *as a non-clash teaching set for* C_1*.*

Definition 2.1.9 *Let* $\mathcal{C}$ *and* $\mathcal{C}'$ *be two concept classes over the same domain* $\mathcal{X}$*. If* $\mathcal{C}$ *is obtained from* $\mathcal{C}'$ *by flipping columns and permuting concepts, then we say* $\mathcal{C}$ *and* $\mathcal{C}'$ *are equivalent. Otherwise, we call them non-equivalent concept classes.*

	x_1	x_2	x_3	x_4	x_5		x_1	x_2	x_3	x_4	x_5
C_1	**1**	0	0	0	**1**	C'_1	**1**	0	**1**	0	1
C_2	**1**	**1**	0	0	0	C'_2	1	**1**	0	**1**	0
C_3	0	**1**	**1**	0	0	C'_3	0	1	**1**	0	**1**
C_4	0	0	**1**	**1**	0	C'_4	**1**	0	1	**1**	0
C_5	0	0	0	**1**	**1**	C'_5	0	**1**	0	1	**1**

Table 2.4: Warmuth's class $\mathcal{C}_W$ with RTD = 3 and VCD = 2.

The following property, applied to both finite and infinite concept classes, expresses a general relation between non-clash teaching parameters and the previously mentioned teaching complexities [24]:

Proposition 2.1.10 *For any concept class* $\mathcal{C}$*, we have* $\mathrm{NCTD}(\mathcal{C}) \leq \mathrm{PBTD}(\mathcal{C}) \leq \mathrm{RTD}(\mathcal{C})$ *and* $\mathrm{NCTD}^+(\mathcal{C}) \leq \mathrm{PBTD}^+(\mathcal{C}) \leq \mathrm{RTD}^+(\mathcal{C})$.

Given the above, there are a small number of concept classes $\mathcal{C}$ which satisfy the inequality $2{\cdot}\mathrm{NCTD}(\mathcal{C}) \leq \mathrm{RTD}(\mathcal{C})$. As an example, the powerset $\mathcal{P}_m$ over the domain $\{x_1, \ldots, x_m\}$, satisfies $\mathrm{NCTD}(\mathcal{P}_m) = \lceil \frac{m}{2} \rceil$ and $\mathrm{NCTD}^+(\mathcal{P}_m) = m$, [24, Theorem 19].

Example 2.1.11 *Warmuth's class: Over the domain* $\mathcal{X} = \{x_1, \ldots, x_5\}$*, consider the 10 concepts shown in Table 2.4. Highlighted labels correspond to a positive non-clashing mapping for* $\mathcal{C}_W$*; it can be shown that* $\mathrm{NCTD}(\mathcal{C}_W) = \mathrm{NCTD}^+(\mathcal{C}_W) = 2$*, [24, Proposition 27].*

This class is also considered as an example for which RTD exceeds VCD. In fact, the following theorem is proved in [7]:

Theorem 2.1.12 *Let $\mathcal{C}$ be a concept class over domain $\mathcal{X}$ such that* RTD > VCD. *Then $|\mathcal{C}| \geq 10$ and $|\mathcal{X}| \geq 5$.*

We continue by providing some further definitions in order to pave the way to obtain a lower bound on NCTD and NCTD$^+$.

The *symmetric difference* of two concepts C and C', denoted by $C \triangle C'$, is defined to be $(C \setminus C') \cup (C' \setminus C)$, and $|C \triangle C'|$ is the number of instances for which C and C' disagree. The symmetric difference between concepts C and C' is also called the *Hamming distance* between the binary vector representations of C and C'. In a concept class $\mathcal{C}$, if two distinct concepts have Hamming distance 1, then they are called *neighbors*. The degree of C, denoted by $\deg_{\mathcal{C}}(C)$, is then defined to be the number of neighbors of C in $\mathcal{C}$. Further we let

$$\deg_{avg}(\mathcal{C}) := \frac{1}{|\mathcal{C}|} \cdot \sum_{C \in \mathcal{C}} \deg_{\mathcal{C}}(C) \tag{2.1}$$

be the average degree of concepts in $\mathcal{C}$.

Using this neighbor notion, the *dominance* of $C \in \mathcal{C}$, denoted as $\mathrm{dom}_{\mathcal{C}}(C)$, is defined as the number of smaller neighbors of C in $\mathcal{C}$, that is neighbors that contain exactly one fewer instance than C.

According to the above definitions, the following lower bounds have been obtained in [24].

Theorem 2.1.13 *Every concept class $\mathcal{C}$ over a finite domain satisfies*

$$\mathrm{NCTD}(\mathcal{C}) \geq \left\lceil \frac{1}{2} \cdot \deg_{avg}(\mathcal{C}) \right\rceil . \tag{2.2}$$

Theorem 2.1.14 *Every concept class $\mathcal{C}$ over a finite domain satisfies*

$$\mathrm{NCTD}^{+}(\mathcal{C}) \geq \max_{C \in \mathcal{C}} \mathrm{dom}_{\mathcal{C}}(C). \tag{2.3}$$

2.2 Graph Notions

We close this section by introducing some definitions from graph theory [6]. They are helpful once a graph is assigned to a concept class to investigate potential properties of its teaching sets.

We use the notation $G = (V, E)$ to denote the simple graph with nonempty vertex set $V = V(G)$ and edge set $E = E(G) \subseteq \{\{u, v\} : u \neq v \in V\}$. The order of the graph G is defined to be $|V|$. For simplicity, one can point out an edge e of G with end points u and v by $e = uv$. The end points u and v are said to be *incident* with the edge $e = uv$. Indicating each vertex of a graph by a single point, and an edge by joining two such vertices, will provide a graphical representation which give us some clues regarding its properties, and results in some new definitions and concepts. Two vertices u, v of G are *neighbors*, if there is and edge between them. The set of neighbors of a vertex v in a graph G is denoted by $N_G(v)$. The *degree* $d_G(v)$ of a vertex v is equal to the number of neighbors of v, i.e. $|N_G(v)|$. The number

$$\deg_{avg}(G) = \frac{1}{|V|}\sum_{v\in V} d_G(v) \tag{2.4}$$

is the *average degree* of graph G.

By summing up all the vertex degrees in G, every edge is counted exactly twice. Thus, one can conclude that $\deg_{avg}(G) = 2\frac{|E|}{|V|}$, which is well-known as the *Handshaking* property.

A graph H is a subgraph of a graph G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. Moreover, a subgraph H of G is said to be induced if for any vertices u and v of H, uv is an edge of H if and only if uv is an edge of G.

A *path* on $n \geq 1$ vertices is a graph P_n with vertex set $V(P_n) = \{v_1, \ldots, v_n\}$ and edge set $E(P_n) = \{v_iv_{i+1} : i = 1, \ldots, n-1\}$. A disjoint union of trees is called a *forest*. A *cycle* of order $n \geq 3$, denoted by C_n, is a graph with vertex set $V(C_n) = \{v_1, \ldots, v_n\}$ and edge set $E(C_n) = \{v_iv_{i+1} : i = 1, \ldots, n-1\} \cup v_1v_n$. The *length* of a path or a cycle is the number of its edges. The distance $d_G(u, v)$ in G of two vertices u, v is the length of a shortest path in G between u and v.

A graph is called *connected* if there is a path between every pair of vertices; otherwise, the graph is *disconnected*.

Chapter 3

On the Ratio of RTD and NCTD

It was formally shown by Kirkpatrick et al. [24] that the non-clash teaching model is optimal among all models satisfying Goldman and Mathias's notion of collusion-freeness, i.e., compared to all other collusion-free teaching models, no-clash teaching requires the smallest worst-case number of examples for teaching a concept in the underlying concept class $\mathcal{C}$. However, so far we are only aware of concept classes for which the ratio between RTD and NCTD is at most 2. The main focus of this chapter is to answer the question whether there exists a concept class for which $\frac{\text{RTD}}{\text{NCTD}} > 2$ holds.

3.1 A Smallest Concept Class with $\text{RTD} = 3 \cdot \text{NCTD}$

To construct a concept class for which the ratio $\frac{\text{RTD}}{\text{NCTD}}$ exceeds 2, we start by considering a concept class containing two concepts C_1 and C_2 over the domain $\{x_1, x_2\}$ in order to build a bigger concept class which satisfies $\mathrm{TD}_{\min}(\mathcal{C}) > 2$. Meanwhile, we try to preserve the no-clash teaching property so that $\mathrm{NCTD}^{+}(\mathcal{C}) = 1$ is satisfied. From these constraints, we will first derive a necessary condition that every concept class with $\frac{\mathrm{RTD}}{\mathrm{NCTD}^{+}} > 2$ fulfills.

Without loss of generality, we are considering concept classes for which a positive non-clash teacher mapping teaches concepts C_i with the example $(x_i, 1)$, $1 \leq i \leq 2$. Since $\mathrm{NCTD}^{+}(\mathcal{C}) = 1$, initially, we have the matrix $\begin{bmatrix} \mathbf{1} & * \\ * & \mathbf{1} \end{bmatrix}$, where each $*$ symbol can be independently replaced by a value in $\in \{0, 1\}$, which results in the three following cases. Observe that since our teacher mapping is non-clashing, the case $\begin{bmatrix} \mathbf{1} & 1 \\ 1 & \mathbf{1} \end{bmatrix}$ does not need to be considered.

Case 1: $\begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{bmatrix}$

By considering the fact that for a concept class $\mathcal{C}$, $\mathrm{TD}_{min}(\mathcal{C}) > 2$ leads to $\mathrm{RTD}(\mathcal{C}) \geq 3$, we will form the following matrices. In each step a new concept and instance are added in order to prevent sets of size two, highlighted by circled labels in each row, from becoming teaching sets. For instance, the fifth row of the matrices below is added to prevent the two circled labels in the second matrix from forming a teaching set. Meanwhile, we also need to replace red stars with 0 or 1 to allow the bold 1's on the diagonal to be non-clashing teaching sets of size one. For example, the red stars in the first two rows of the first matrix below must be replaced by 0 to avoid clashes with the teaching sets for the third and fourth rows.

$$
\begin{bmatrix} \textcircled{1} & \textcircled{0} & {\color{red}*} & * \\ \textcircled{0} & \textcircled{1} & * & {\color{red}*} \\ 1 & 0 & \mathbf{1} & * \\ 0 & 1 & * & \mathbf{1} \end{bmatrix}
\xrightarrow[\text{clash}]{\text{no}}
\begin{bmatrix} \mathbf{1} & 0 & 0 & * \\ 0 & \mathbf{1} & * & 0 \\ \textcircled{1} & 0 & \textcircled{1} & * \\ 0 & 1 & * & \mathbf{1} \end{bmatrix}
\rightarrow
\begin{bmatrix} \mathbf{1} & 0 & 0 & * & {\color{red}*} \\ 0 & \mathbf{1} & * & 0 & * \\ 1 & 0 & \mathbf{1} & * & {\color{red}*} \\ 0 & 1 & * & \mathbf{1} & * \\ 1 & * & 1 & * & \mathbf{1} \end{bmatrix}
\xrightarrow[\text{clash}]{\text{no}}
\begin{bmatrix} \mathbf{1} & 0 & 0 & * & 0 \\ 0 & \mathbf{1} & * & 0 & * \\ 1 & 0 & \mathbf{1} & * & 0 \\ \textcircled{0} & 1 & * & \textcircled{1} & * \\ 1 & * & 1 & * & \mathbf{1} \end{bmatrix}
\rightarrow
\begin{bmatrix} \mathbf{1} & 0 & 0 & * & 0 & * \\ 0 & \mathbf{1} & * & 0 & * & * \\ 1 & 0 & \mathbf{1} & * & 0 & * \\ 0 & 1 & * & \mathbf{1} & * & {\color{red}*} \\ 1 & * & 1 & * & \mathbf{1} & * \\ 0 & * & * & 1 & * & \mathbf{1} \end{bmatrix}
$$

$$
\xrightarrow[\text{clash}]{\text{no}}
\begin{bmatrix} \mathbf{1} & 0 & 0 & * & 0 & * \\ \textcircled{0} & \mathbf{1} & * & \textcircled{0} & * & * \\ 1 & 0 & \mathbf{1} & * & 0 & * \\ 0 & 1 & * & \mathbf{1} & * & 0 \\ 1 & * & 1 & * & \mathbf{1} & * \\ 0 & * & * & 1 & * & \mathbf{1} \end{bmatrix}
\rightarrow
\begin{bmatrix} \textcircled{1} & 0 & \textcircled{0} & * & 0 & * & * \\ 0 & \mathbf{1} & * & 0 & * & * & * \\ 1 & 0 & \mathbf{1} & * & 0 & * & * \\ 0 & 1 & * & \mathbf{1} & * & 0 & * \\ 1 & * & 1 & * & \mathbf{1} & * & * \\ 0 & * & * & 1 & * & \mathbf{1} & * \\ 0 & * & * & 0 & * & * & \mathbf{1} \end{bmatrix}
\rightarrow
\begin{bmatrix} \mathbf{1} & 0 & 0 & * & 0 & * & * & {\color{red}*} \\ 0 & \mathbf{1} & * & 0 & * & * & * & * \\ 1 & 0 & \mathbf{1} & * & 0 & * & * & * \\ 0 & 1 & * & \mathbf{1} & * & 0 & * & * \\ 1 & * & 1 & * & \mathbf{1} & * & * & * \\ 0 & * & * & 1 & * & \mathbf{1} & * & * \\ 0 & * & * & 0 & * & * & \mathbf{1} & * \\ 1 & * & 0 & * & * & * & * & \mathbf{1} \end{bmatrix}
$$

$$
\xrightarrow[\text{clash}]{\text{no}}
\begin{bmatrix} \mathbf{1} & 0 & 0 & * & 0 & * & * & 0 \\ 0 & \mathbf{1} & * & 0 & * & * & * & * \\ 1 & 0 & \mathbf{1} & * & 0 & * & * & * \\ 0 & 1 & * & \mathbf{1} & * & 0 & * & * \\ 1 & * & 1 & * & \mathbf{1} & * & * & * \\ 0 & * & * & 1 & * & \mathbf{1} & * & * \\ 0 & * & * & 0 & * & * & \mathbf{1} & * \\ 1 & * & 0 & * & * & * & * & \mathbf{1} \end{bmatrix}.
$$

Case 2: $\begin{bmatrix} \mathbf{1} & 1 \\ 0 & \mathbf{1} \end{bmatrix}$

$$
\begin{bmatrix} \textcircled{1} & \textcircled{1} & {\color{red}*} & * \\ \textcircled{0} & \textcircled{1} & {\color{red}*} & {\color{red}*} \\ 1 & 1 & \mathbf{1} & * \\ 0 & 1 & * & \mathbf{1} \end{bmatrix}
\xrightarrow[\text{clash}]{\text{no}}
\begin{bmatrix} \mathbf{1} & 1 & 0 & * \\ 0 & \mathbf{1} & 0 & 0 \\ \textcircled{1} & 1 & \textcircled{1} & * \\ 0 & 1 & * & \mathbf{1} \end{bmatrix}
\rightarrow
\begin{bmatrix} \mathbf{1} & 1 & 0 & * & {\color{red}*} \\ 0 & \mathbf{1} & 0 & 0 & * \\ 1 & 1 & \mathbf{1} & * & {\color{red}*} \\ 0 & 1 & * & \mathbf{1} & * \\ 1 & * & 1 & * & \mathbf{1} \end{bmatrix}
\xrightarrow[\text{clash}]{\text{no}}
\begin{bmatrix} \mathbf{1} & 1 & 0 & * & 0 \\ 0 & \mathbf{1} & 0 & 0 & * \\ 1 & 1 & \mathbf{1} & * & 0 \\ \textcircled{0} & 1 & * & \textcircled{1} & * \\ 1 & * & 1 & * & \mathbf{1} \end{bmatrix}
\rightarrow
\begin{bmatrix} \mathbf{1} & 1 & 0 & * & 0 & * \\ 0 & \mathbf{1} & 0 & 0 & * & * \\ 1 & 1 & \mathbf{1} & * & 0 & * \\ 0 & 1 & * & \mathbf{1} & * & {\color{red}*} \\ 1 & * & 1 & * & \mathbf{1} & * \\ 0 & * & * & 1 & * & \mathbf{1} \end{bmatrix}
$$

$$
\xrightarrow[\text{clash}]{\text{no}}
\begin{bmatrix} \mathbf{1} & 1 & 0 & * & 0 & * \\ \textcircled{0} & \mathbf{1} & 0 & \textcircled{0} & * & * \\ 1 & 1 & \mathbf{1} & * & 0 & * \\ 0 & 1 & * & \mathbf{1} & * & 0 \\ 1 & * & 1 & * & \mathbf{1} & * \\ 0 & * & * & 1 & * & \mathbf{1} \end{bmatrix}
\rightarrow
\begin{bmatrix} \textcircled{1} & 1 & \textcircled{0} & * & 0 & * & * \\ 0 & \mathbf{1} & 0 & 0 & * & * & * \\ 1 & 1 & \mathbf{1} & * & 0 & * & * \\ 0 & 1 & * & \mathbf{1} & * & 0 & * \\ 1 & * & 1 & * & \mathbf{1} & * & * \\ 0 & * & * & 1 & * & \mathbf{1} & * \\ 0 & * & * & 0 & * & * & \mathbf{1} \end{bmatrix}
\rightarrow
\begin{bmatrix} \mathbf{1} & 1 & 0 & * & 0 & * & * & {\color{red}*} \\ 0 & \mathbf{1} & 0 & 0 & * & * & * & * \\ 1 & 1 & \mathbf{1} & * & 0 & * & * & * \\ 0 & 1 & * & \mathbf{1} & * & 0 & * & * \\ 1 & * & 1 & * & \mathbf{1} & * & * & * \\ 0 & * & * & 1 & * & \mathbf{1} & * & * \\ 0 & * & * & 0 & * & * & \mathbf{1} & * \\ 1 & * & 0 & * & * & * & * & \mathbf{1} \end{bmatrix}
$$

$$\xrightarrow[\text{clash}]{\text{no}} \begin{bmatrix} \mathbf{1} & 1 & 0 & \ast & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & \ast & \ast & \ast & \ast \\ 1 & 1 & \mathbf{1} & \ast & 0 & \ast & \ast & \ast \\ 0 & 1 & \ast & \mathbf{1} & \ast & 0 & \ast & \ast \\ 1 & \ast & 1 & \ast & \mathbf{1} & \ast & \ast & \ast \\ 0 & \ast & \ast & 1 & \ast & \mathbf{1} & \ast & \ast \\ 0 & \ast & \ast & 0 & \ast & \ast & \mathbf{1} & \ast \\ 1 & \ast & 0 & \ast & \ast & \ast & \ast & \mathbf{1} \end{bmatrix}.$$

Case 3: $\begin{bmatrix} \mathbf{1} & 0 \\ 1 & \mathbf{1} \end{bmatrix}$ generates the same 8 by 8 matrix as obtained in Case 2.

According to the above cases, the following result is the first fact that can be established.

Theorem 3.1.1 *Over a domain $\mathcal{X}$, if a concept class $\mathcal{C}$ satisfies the equalities* $\mathrm{RTD}(\mathcal{C}) = 3$ *and* $\mathrm{NCTD}^+(\mathcal{C}) = 1$*, then both $\mathcal{X}$ and $\mathcal{C}$ have at least 8 elements.*

The next step is to replace the stars in the final matrices above by actual labels (0 or 1) in order to construct matrices for which RTD equals 3 and NCTD^+ equals 1. Working with the matrix provided in case 2 leads to our desired result, if we choose to replace the four stars in the second column with 0's. The steps of forming a new concept class are illustrated as follows. Circles and red stars in this illustration play the same role as before; blue stars indicate where a label is chosen to avoid a teaching set of size two.

$$\begin{bmatrix} \mathbf{1} & 1 & 0 & \ast & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & \ast & \ast & \ast & \ast \\ 1 & \textcircled{1} & \textcircled{\textbf{1}} & \ast & 0 & \ast & \ast & \ast \\ 0 & 1 & \textcolor{blue}{\ast} & \mathbf{1} & \ast & 0 & \ast & \ast \\ 1 & 0 & 1 & \ast & \mathbf{1} & \ast & \ast & \ast \\ 0 & 0 & \ast & 1 & \ast & \mathbf{1} & \ast & \ast \\ 0 & 0 & \ast & 0 & \ast & \ast & \mathbf{1} & \ast \\ 1 & 0 & 0 & \ast & \ast & \ast & \ast & \mathbf{1} \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{1} & 1 & 0 & \ast & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & \ast & \ast & \ast & \ast \\ 1 & 1 & \mathbf{1} & \textcolor{red}{\ast} & 0 & \ast & \ast & \ast \\ 0 & 1 & \textcolor{blue}{1} & \mathbf{1} & \ast & 0 & \ast & \ast \\ 1 & 0 & 1 & \ast & \mathbf{1} & \ast & \ast & \ast \\ 0 & 0 & \ast & 1 & \ast & \mathbf{1} & \ast & \ast \\ 0 & 0 & \ast & 0 & \ast & \ast & \mathbf{1} & \ast \\ 1 & 0 & 0 & \ast & \ast & \ast & \ast & \mathbf{1} \end{bmatrix} \xrightarrow[\text{clash}]{\text{no}} \begin{bmatrix} \mathbf{1} & 1 & 0 & \textcolor{blue}{\ast} & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & \ast & \ast & \ast & \ast \\ 1 & 1 & \mathbf{1} & \textcolor{red}{0} & 0 & \ast & \ast & \ast \\ 0 & \textcircled{1} & \textcolor{blue}{1} & \textcircled{\textbf{1}} & \ast & 0 & \ast & \ast \\ 1 & 0 & 1 & \ast & \mathbf{1} & \ast & \ast & \ast \\ 0 & 0 & \ast & 1 & \ast & \mathbf{1} & \ast & \ast \\ 0 & 0 & \ast & 0 & \ast & \ast & \mathbf{1} & \ast \\ 1 & 0 & 0 & \ast & \ast & \ast & \ast & \mathbf{1} \end{bmatrix}$$

$$\to \begin{bmatrix} \mathbf{1} & 1 & \textcircled{0} & \textcircled{1} & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & \textcircled{0} & \textcircled{0} & \ast & \ast & \ast & \ast \\ 1 & 1 & \textcircled{\textbf{1}} & \textcircled{0} & 0 & \ast & \ast & \ast \\ 0 & 1 & \textcircled{1} & \textcircled{\textbf{1}} & \ast & 0 & \ast & \ast \\ 1 & 0 & 1 & \ast & \mathbf{1} & \ast & \ast & \ast \\ 0 & 0 & \ast & 1 & \ast & \mathbf{1} & \ast & \ast \\ 0 & 0 & \ast & 0 & \ast & \ast & \mathbf{1} & \ast \\ 1 & 0 & 0 & \ast & \ast & \ast & \ast & \mathbf{1} \end{bmatrix} \to \begin{bmatrix} \mathbf{1} & 1 & 0 & 1 & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & \ast & \ast & \ast & \ast \\ 1 & 1 & \mathbf{1} & 0 & 0 & \ast & \ast & \ast \\ 0 & 1 & 1 & \mathbf{1} & \ast & 0 & \ast & \ast \\ 1 & 0 & 1 & \textcircled{1} & \textcircled{\textbf{1}} & \ast & \ast & \ast \\ 0 & 0 & 0 & 1 & \ast & \mathbf{1} & \ast & \ast \\ 0 & 0 & 1 & 0 & \ast & \ast & \mathbf{1} & \ast \\ 1 & 0 & 0 & 0 & \ast & \ast & \ast & \mathbf{1} \end{bmatrix} \xrightarrow[\text{clash}]{\text{no}} \begin{bmatrix} \mathbf{1} & 1 & 0 & 1 & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & \ast & \ast & \ast & \ast \\ 1 & 1 & \mathbf{1} & 0 & 0 & \ast & 0 & \ast \\ 0 & 1 & 1 & \mathbf{1} & 0 & 0 & \ast & \ast \\ 1 & 0 & \textcircled{\textbf{1}} & 1 & \textcircled{\textbf{1}} & \ast & \ast & \ast \\ 0 & 0 & 0 & 1 & 1 & \mathbf{1} & \ast & \ast \\ 0 & 0 & 1 & 0 & \ast & \ast & \mathbf{1} & \ast \\ 1 & 0 & 0 & 0 & \ast & \ast & \ast & \mathbf{1} \end{bmatrix}$$

$$\xrightarrow[\text{clash}]{\text{no}} \begin{bmatrix} \mathbf{1} & 1 & 0 & 1 & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & \bigcirc & \ast & \ast & \ast \\ 1 & 1 & \mathbf{1} & 0 & 0 & \ast & 0 & \ast \\ 0 & 1 & 1 & \mathbf{1} & 0 & 0 & \ast & \ast \\ 1 & 0 & 1 & 1 & \mathbf{1} & 0 & 0 & \ast \\ 0 & 0 & 0 & 1 & 1 & \mathbf{1} & \ast & \ast \\ 0 & 0 & 1 & 0 & 1 & \ast & \mathbf{1} & \ast \\ 1 & 0 & 0 & 0 & \bigcirc & \ast & \ast & \mathbf{1} \end{bmatrix}.$$

In the last matrix above the upper and lower circles must be replaced by 0 and 1, respectively. Otherwise we will have a teaching set of size two. Thus

$$\begin{bmatrix} \mathbf{1} & 1 & 0 & 1 & 0 & \ast & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & \ast & \ast & \ast \\ 1 & 1 & \mathbf{1} & 0 & 0 & \ast & 0 & \ast \\ 0 & 1 & 1 & \mathbf{1} & 0 & 0 & \ast & \ast \\ 1 & 0 & 1 & 1 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \textcircled{1} & 1 & \textcircled{\textbf{1}} & \ast & \ast \\ 0 & 0 & 1 & 0 & 1 & \ast & \mathbf{1} & \ast \\ \textcircled{1} & 0 & 0 & 0 & 1 & \ast & \ast & \textcircled{\textbf{1}} \end{bmatrix} \to \begin{bmatrix} \mathbf{1} & 1 & 0 & 1 & 0 & 1 & \ast & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & \ast & \ast & \ast \\ 1 & 1 & \mathbf{1} & 0 & 0 & \ast & 0 & 1 \\ 0 & 1 & 1 & \mathbf{1} & 0 & 0 & \ast & \ast \\ 1 & 0 & 1 & 1 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & \mathbf{1} & \ast & \ast \\ 0 & 0 & 1 & 0 & 1 & \ast & \mathbf{1} & \ast \\ 1 & 0 & 0 & 0 & 1 & \ast & \ast & \mathbf{1} \end{bmatrix}.$$

Continuing the process and replacing the stars by actual labels, we will be able to form the concept classes $\mathcal{C}_S$ and $\mathcal{C'}_S$, shown in Table 3.1, for which RTD $= 3$ and NCTD$^+ = 1$. The bold 1's in Table 3.1 correspond to the positive no-clash teacher mappings for $\mathcal{C}_S$ and $\mathcal{C'}_S$. Moreover, it is easy to see that VCD $= 3$ for both concept classes.

The above discussion leads to the following observation, which is one of the main

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
C_1	**1**	1	0	1	0	1	1	0
C_2	0	**1**	0	0	0	0	0	0
C_3	1	1	**1**	0	0	1	0	1
C_4	0	1	1	**1**	0	0	1	1
C_5	1	0	1	1	**1**	0	0	0
C_6	0	0	0	1	1	**1**	0	1
C_7	0	0	1	0	1	1	**1**	0
C_8	1	0	0	0	1	0	1	**1**

$\mathcal{C}_S$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
C_1	**1**	1	0	1	0	1	0	0
C_2	0	**1**	0	0	0	1	1	1
C_3	1	1	**1**	0	0	0	0	1
C_4	0	1	1	**1**	0	0	1	0
C_5	1	0	1	1	**1**	0	0	0
C_6	0	0	0	1	1	**1**	1	0
C_7	0	0	1	0	1	0	**1**	1
C_8	1	0	0	0	1	1	0	**1**

$\mathcal{C}'_S$

Table 3.1: The concept classes $\mathcal{C}_S$ and $\mathcal{C}'_S$ with RTD = 3 and NCTD$^+$ = 1.

results of this chapter.

Theorem 3.1.2 *There is a concept class $\mathcal{C}$ for which* $\text{RTD}(\mathcal{C}) \geq 3 \cdot \text{NCTD}(\mathcal{C})$.

Proof By Theorem 3.1.1, the concept class $\mathcal{C}_S$ (or $\mathcal{C}'_S$) is a concept class for which RTD = 3 and NCTD$^+$ = 1. Since for any concept class $\mathcal{C}$, $\text{NCTD}^+(\mathcal{C}) \geq \text{NCTD}(\mathcal{C})$, we have the desired result. □

By removing the last four instances of the domain of the concept class $\mathcal{C}_S$, and by allowing negative labels in the teaching sets, we will be able to form a concept class $\mathcal{C}_{3/1}$, provided in Table 3.2, which satisfies the inequality RTD $\geq 3 \cdot$ NCTD. On the other hand, if a concept class has seven or fewer concepts, then its VC-dimension is at most 2, and a simple argument will show that its RTD is at most 2 as well. Therefore, $\mathcal{C}_{3/1}$ is the smallest concept class with RTD $\geq 3 \cdot$ NCTD.

In fact, similar to the 8 by 8 case, one could start from the matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and build the following chain of embedded matrices step by step:

	x_1	x_2	x_3	x_4
C_1	**1**	1	0	1
C_2	0	**1**	0	0
C_3	1	1	**1**	0
C_4	0	1	1	**1**
C_5	1	**0**	1	1
C_6	0	0	**0**	1
C_7	**0**	0	1	0
C_8	1	0	0	**0**

Table 3.2: The concept class $\mathcal{C}_{3/1}$ with RTD = VCD = 3 and NCTD = 1.

	x_1	x_2	x_3	x_4
C_1	**1**	1	**0**	1
C_2	**0**	**1**	0	0
C_3	**1**	1	**1**	0
C_4	**0**	1	**1**	1
C_5	**1**	1	1	**1**
C_6	**0**	1	0	**1**
C_7	0	1	**1**	**0**
C_8	**1**	1	0	**0**

Table 3.3: The concept class $\mathcal{C}_{3/2}$ with RTD = 3 and NCTD = VCD = 2.

$$\begin{bmatrix} 1 & 1 & * & * \\ 0 & 1 & * & * \\ 1 & 1 & 1 & * \\ 0 & 1 & * & 1 \end{bmatrix} \xrightarrow{no-clash} \begin{bmatrix} 1 & 1 & 0 & * \\ \textcircled{0} & 1 & \textcircled{0} & 0 \\ \textcircled{1} & 1 & \textcircled{1} & * \\ 0 & 1 & * & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \textcircled{1} & 1 & \textcircled{0} & * \\ \textcircled{0} & 1 & 0 & \textcircled{0} \\ 1 & 1 & 1 & * \\ 0 & 1 & * & 1 \\ 1 & * & 1 & * \\ 0 & * & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 & * \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & * \\ 0 & 1 & * & 1 \\ 1 & * & 1 & * \\ 0 & * & 0 & 1 \\ 0 & * & * & 0 \\ 1 & * & 0 & * \end{bmatrix}$$

The reader can observe that replacing stars randomly with other labels, may result in other concept classes such as $\mathcal{C}_{3/2}$, shown in Table 3.3, or $\mathcal{C}_{2/2}$, shown in Table 3.4, in which the concept C_7 is removed as it is identical to the concept C_2.

	x_1	x_2	x_3	x_4
C_1	**1**	**1**	0	1
C_2	**0**	**1**	0	0
C_3	**1**	1	**1**	0
C_4	**0**	1	0	**1**
C_5	1	1	**1**	**1**
C_6	**0**	**0**	0	1
C_8	**1**	1	0	**0**

Table 3.4: The concept class $\mathcal{C}_{2/2}$ with RTD = NCTD = VCD = 2.

3.2 Concept Classes Resulting from Block Designs

Research in design theory is of relevance for practical applications in coding theory and for advanced research in algebra, number theory, statistics and geometry. The origins of design theory can be traced back to the statistical works of Fisher and Yates [10], who were inspired by questions of the design of field experiments in agriculture, and showed using experimental analysis how several types of plants can be compared. During the past decades, this area of investigation has introduced numerous new applications and many researchers have been interested in working on specific combinatorial designs, such as balanced incomplete block designs, Hadamard matrices, orthogonal arrays, etc., which has led to powerful new combinatorial and computational techniques [33]. There exist a variety of studied designs that were once considered mathematical puzzles or riddles. For example, Kirkman's schoolgirl problem that goes back to 1850 is one of them. It says "Fifteen young ladies in a school walk out three abreast for seven days in succession: it is required to arrange them daily, so that no two shall walk twice abreast." A 2-$(15, 3, 1)$-design, which is defined later, is a solution to this problem [23].

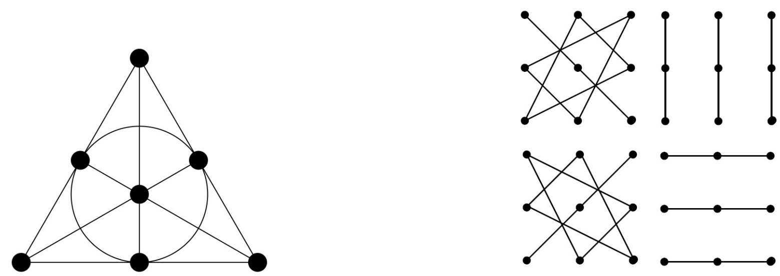
The advent of design theory also resulted in finding connections with certain branches of algebra and geometry. Geometric structures called "finite projective planes", collections of points and lines which satisfy some mathematical properties, can be considered as a well-known structure in these branches. Such connections provide mathematicians a geometric way to introduce particular designs [13, 22]. Designs have many other applications as well, including, but not limited to cryptography, coding theory, mathematical biology, algorithm design and analysis, and networking. The following is the definition of a t-design:

Definition 3.2.1 *A t-(v, k, λ) design $\mathcal{D} = (X, \mathcal{B})$ consists of a set X of size v and a collection $\mathcal{B}$ of k-subsets*[1] *of X, with the property that every t-subset of X is contained in exactly λ members of $\mathcal{B}$. The elements of X are called points and the elements of $\mathcal{B}$ are called blocks. A t-design is nontrivial if $0 < t < k < v$ and not every k-subset of X is a block.*

The following are some basic examples:

Example 3.2.2 *A* 2-$(7, 3, 1)$ *design:* $X = \{1, 2, 3, 4, 5, 6, 7\}$ *and* $\mathcal{B} = \{123, 145, 167, 246, 257, 347, 356\}$. *This design is the smallest projective plane and is a collection*

[1] A k-subset is a subset of X that contains exactly k elements.



(a) The Fano Plane. (b) A 2-$(9, 3, 1)$ design.

Figure 3.1: The Fano Plane and a 2-$(9, 3, 1)$ design.

of seven points which is known as the Fano plane. Each line contains exactly three points, and each pair of points appears in exactly one line, see Figure 3.1(a).

Example 3.2.3 *A* 2-$(9, 3, 1)$ *design:* $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ *and* $\mathcal{B} = \{123, 456, 789, 147, 258, 369, 159, 267, 348, 168, 249, 357\}$. *The graph of this design contains eight lines and four triangles, and is presented in Figure 3.1(b).*

Example 3.2.4 *A* 3-$(5, 3, 3)$ *design:* $X = \{1, 2, 3, 4, 5\}$ *and* $\mathcal{B} = \{123, 124, 125, 134, 135, 145, 234, 235, 245, 345\}$.

Based on the three fixed parameters (v, k, λ), we can determine two other parameters, labelled r and b, which are consequently fixed. Therefore, one sometimes finds the notation 2-(v, b, r, k, λ) design in the literature to refer to a 2-(v, k, λ) design. We have the following fact [30, 33]:

Theorem 3.2.5 *In a* 2-(v, k, λ) *design* $(X, \mathcal{B})$*, every point occurs in exactly* $r = \frac{\lambda(v-1)}{k-1}$ *blocks and* $\mathcal{B}$ *has exactly* $b = \frac{vr}{k}$ *blocks.*

Since b and r must be integers, 2-designs with certain parameters might not exist. For example, 2-$(8,3,1)$ and 2-$(19,4,1)$ designs do not exist [33]. Therefore, determining necessary and sufficient conditions for the existence or non-existence of a 2-(v,k,λ) design can be considered as one of the fundamental goals of combinatorial design theory. In addition, the equality $v = b$, equivalently $k = r$, defines a special kind of design called a *symmetric design* which is often used in codes. Following are two theorems in which all these parameters play a role [33].

Theorem 3.2.6 *If a* 2-(v,k,λ) *design exists, then* $\lambda(v-1) \equiv 0 \ (mod\ k-1)$ *and* $\lambda v(v-1) \equiv 0 \ (mod\ k(k-1))$.

Theorem 3.2.7 *If there exists a 2-design with parameters* (v,b,r,k,λ)*, then* $vr = bk$, $r(k-1) = \lambda(v-1)$ *and* $b > r > \lambda$ *must hold.*

A representation of a design $\mathcal{D} = (X, \mathcal{B})$ is a matrix in which there is a row for each instance v of X and the entries in this row correspond to the block elements. Here is an exact definition of the incidence matrix of a design [30].

Definition 3.2.8 *Let* $(X, \mathcal{B})$ *a design where* $X = \{x_1, \ldots, x_v\}$ *and* $\mathcal{B} = \{B_1, \ldots, B_b\}$*. The incidence matrix of* $(X, \mathcal{B})$ *is the matrix* $M = [m_{ij}]_{v \times b}$ *defined by the rule*

$$m_{ij} = \begin{cases} 1 & \text{if } x_i \in B_j \\ 0 & \text{if } x_i \notin B_j. \end{cases}$$

Example 3.2.9 *Consider the* 2-$(7,3,1)$ *and* 2-$(5,3,3)$ *designs presented in Examples 3.2.2 and 3.2.4. The incidence matrices of these designs are the following* 7×7 *and* 5×10 *matrices, respectively.*

$$M = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}_{7\times 7}, \quad M = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}_{5\times 10}.$$

During the past few years, new ideas in constructing designs have been proposed and new combinatorial techniques have been found in order to construct families of designs. An eager reader may have a look at references [5] and [30] to investigate other types of block designs, such as group divisible designs and cyclic designs.

Large families of 2- and 3-designs can be constructed using *Hadamard matrices.* These matrices have many applications and have been investigated since the 1860s. For more details see [22].

Definition 3.2.10 *A Hadamard matrix of order* n *is an* $n \times n$ *matrix* H_n *with entries* ± 1 *which satisfies* $H_n H_n^T = nI_n$*, where* I_n *is the* $n \times n$ *identity matrix.*

Hadamard found that a necessary condition for the existence of such matrix is that $n \in \{1, 2\}$, or n is a positive multiple of 4. It is not known whether a Hadamard matrix exists for every multiple of 4 but a very large quantity of them exist, including some well-known infinite families of such matrices. By adding a full row of 1s and a full column of 1s to the incidence matrix of the 2-$(7, 3, 1)$ design in Example 3.2.9, a Hadamard matrix of order 8 is obtained. This fact comes from a general result which shows a natural link between Hadamard matrices and 2-designs [33].

Theorem 3.2.11 *Let n be a positive integer and let H be a ± 1-matrix of order $4n$ with all entries in the first row and the first column equal to 1. Let A be the matrix of order $4n-1$ obtained by removing the first row and the first column of H and let $N = \frac{1}{2}(A+J)$. Then H is a Hadamard matrix if and only if N is an incidence matrix of a symmetric 2-$(4n-1, 2n-1, n-1)$ design.*

The symmetric designs in the above theorem are called *Hadamard 2-designs of order n*. The following result provides another design associated to a Hadamard matrix [33].

Theorem 3.2.12 *Let n be a positive integer and let $H = [a_{ij}]$ be a Hadamard matrix of order $4n$ with all entries in the last row equal to 1. Let $X = \{1, 2, \ldots, 4n\}$. For $i = 1, 2, \ldots, 4n-1$, let $A_i = \{j \in X : a_{ij} = 1\}$ and $B_i = \{j \in X : a_{ij} = -1\}$. Then the incidence structure $\mathcal{D} = (X, \mathcal{B})$ where $\mathcal{B} = \{A_1, A_2, \ldots, A_{4n-1}, B_1, B_2, \ldots, B_{4n-1}\}$ is a 2-$(4n, 2n, 2n-1)$ design. Furthermore, any 3-subset of X is contained in exactly $n-1$ blocks of $\mathcal{D}$.*

In what follows, we will see that design theory may provide us with some results of interest in research on computational learning theory.

	x_1	x_2	x_3	x_4
C_1	**1**	1	1	1
C_2	1	0	1	**0**
C_3	1	1	**0**	0
C_4	1	**0**	0	1
C_5	**0**	0	0	0
C_6	0	1	0	**1**
C_7	0	0	**1**	1
C_8	0	**1**	1	0

Table 3.5: The concept class $\mathcal{C}'_{3/1}$ with RTD = VCD = 3 and NCTD = 1.

Observation 1: As we mentioned earlier, adding a full row of 1's to the incidence matrix of the 2-$(7,3,1)$ design shown in Example 3.2.9, and then removing some columns, results in the following concept class for which the equalities RTD = 3 and NCTD = 1 hold, as shown in Table 3.5. Observe that flipping 0's for 1's (and vice versa) in only one column of $\mathcal{C}'_{3/1}$ and then permuting columns and permuting rows leads to the same concept class $\mathcal{C}_{3/1}$ shown already in Table 3.2.

Observation 2: Set $\mathcal{C}_0 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, and let its complement be $\widehat{\mathcal{C}_0} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Proceed in the same manner to construct $\mathcal{C}_1$, $\widehat{\mathcal{C}_1}$ and $\mathcal{C}_2$ as follows:

$$\mathcal{C}_1 = \begin{bmatrix} \mathcal{C}_0 & \mathcal{C}_0 \\ \mathcal{C}_0 & \widehat{\mathcal{C}_0} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix}, \qquad \mathcal{C}_2 = \begin{bmatrix} \mathcal{C}_1 & \mathcal{C}_1 \\ \mathcal{C}_1 & \widehat{\mathcal{C}_1} \end{bmatrix} = \begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&0&1&0&1&0&1&0 \\ 1&1&0&0&1&1&0&0 \\ 1&0&0&1&1&0&0&1 \\ 1&1&1&1&0&0&0&0 \\ 1&0&1&0&0&1&0&1 \\ 1&1&0&0&0&0&1&1 \\ 1&0&0&1&0&1&1&0 \end{bmatrix}.$$

It can be seen that $\mathrm{RTD}(\mathcal{C}_2) = 3$ and $\mathrm{NCTD}(\mathcal{C}_2) = 1$. Note that $\mathcal{C}_2$ can be built by rearranging the columns of the incidence matrix of the 2-$(7,3,1)$ design from Example

$$
\mathcal{C}'_3 = \begin{bmatrix} \mathcal{C}_2 \\ \widehat{\mathcal{C}_2} \end{bmatrix} = \begin{bmatrix}
\mathbf{1} & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & \mathbf{0} & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & \mathbf{0} & 1 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & \mathbf{0} & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & \mathbf{0} \\
1 & 0 & 1 & 0 & \mathbf{0} & 1 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & \mathbf{0} & 1 & 1 \\
1 & 0 & \mathbf{0} & 1 & 0 & 1 & 1 & 0 \\
\mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \mathbf{1} & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & \mathbf{1} & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 & \mathbf{1} & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & \mathbf{1} \\
0 & 1 & 0 & 1 & \mathbf{1} & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & \mathbf{1} & 0 & 0 \\
0 & 1 & \mathbf{1} & 0 & 1 & 0 & 0 & 1
\end{bmatrix}.
$$

Table 3.6: The concept class $\mathcal{C}'_3$ with RTD = VCD = 4 and NCTD = 1.

3.2.9 and adding a full row of 1's as well as a full columns of 1's to it.

Observation 3: There exists a concept class $\mathcal{C}$ with NCTD($\mathcal{C}$) = 1 and RTD($\mathcal{C}$) = 4.

By continuing the above process and considering the right half of the matrix $\mathcal{C}_3$ we have the same result for $\mathcal{C}'_3$, see Table 3.6, as witnessed by the entries highlighted in bold corresponding to non-clashing teaching sets, which is a concept class that includes 16 concepts over a domain of size 8.

The process mentioned in Observation 2, when replacing 0 by -1, is identical to forming a family of Hadamard matrices using the operation known as the Kronecker product [26], defined as follows:

Definition 3.2.13 *The Kronecker product of the matrix $A_{m\times n}$ with the matrix $B_{p\times q}$ is the $mp \times nq$-matrix defined as*

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1}B & a_{22}B & \cdots & a_{mn}B \end{bmatrix}.$$

As we mentioned earlier in this section, one open problem in design theory is to decide whether designs with certain parameter values even exist. Hence, there is a limited number of incidence matrices known to be associated with designs. However, we can construct Hadamard matrices via tensor products, which is very useful for our purposes so we focused our investigation on $(0,1)$-matrices with a Hadamard structure. This kind of 2-design structure has been the most intensely studied type historically due to its applications in the design of experiments.

The so-called standard Hadamard matrices of dimension 2^n for $n \in \mathbb{N}$ are given by the recursive formula:

$$H_1 = \begin{bmatrix} 1 \end{bmatrix},$$
$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$
$$H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix}, \text{ and}$$
$$H_{2^n} = \begin{bmatrix} H_{2^{n-1}} & H_{2^{n-1}} \\ H_{2^{n-1}} & -H_{2^{n-1}} \end{bmatrix} = H_2 \otimes H_{2^{n-1}}.$$

Generally, if H_n and H_m are Hadamard matrices of order n and m respectively then their Kronecker product $H_{nm} = H_n \otimes H_m$ is a Hadamard matrix of order nm [26]. As an example, consider the following Hadamard matrices H_1 and H_2:

$$H_1 = \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}.$$

By replacing the -1 entries of $H_1 \otimes H_2$ with zeros, we obtain the following incidence matrix associated with a 2-$(16, 6, 2)$ design:

$$\mathcal{C}_{2-(16,6,2)} = \begin{bmatrix}
\mathbf{0} & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\mathbf{1} & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & \mathbf{1} & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & \mathbf{0} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & \mathbf{1} & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & \mathbf{1} & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & \mathbf{1} & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & \mathbf{1} & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & \mathbf{1} & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \mathbf{1} \\
1 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0
\end{bmatrix}.$$

One can verify that $\mathrm{RTD}(\mathcal{C}_{2\text{-}(16,6,2)}) = 3$, $\mathrm{NCTD}(\mathcal{C}_{2\text{-}(16,6,2)}) = 1$.

Observation 4: As a corollary of Theorem 3.2.12, if there exists a Hadamard matrix of order $4n$, then there exists a 3-design over $X = \{1, 2, \ldots, 4n\}$. Therefore, since both $\mathcal{C}_2$, in Observation 3.2, and $\mathcal{C}_{2\text{-}(16,6,2)}$ are Hadamard matrices, we conclude that both values $\mathrm{NCTD}^+(\mathcal{C}_2)$ and $\mathrm{NCTD}^+(\mathcal{C}_{2\text{-}(16,6,2)})$ are less than or equal to 3.

Proposition 3.2.14 *Let $(X, \mathcal{B})$ be a t-$(v, k, 1)$ design where $v \neq k$. Then*

$\mathrm{TD}^+(\mathcal{C}_{t\text{-}(v,k,1)}) = \mathrm{RTD}^+(\mathcal{C}_{t\text{-}(v,k,1)}) = t$. *Further* $\mathrm{NCTD}^+(\mathcal{C}_{t\text{-}(v,k,1)}) \le t$.

Proof According to Corollary 9.6 in [33], if $(X, \mathcal{B})$ is a t-(v, k, λ) design, and $1 \le s < t$, then $(X, \mathcal{B})$ is a s-(v, k, λ_s) design where $\lambda_s = \lambda \cdot \frac{\binom{v-s}{t-s}}{\binom{k-s}{t-s}}$. Since $(X, \mathcal{B})$ is a t-$(v, k, 1)$ design, every t-subset of X is contained in exactly $\lambda = 1$ member of $\mathcal{B}$. On the other hand, according to our assumption $\lambda_s > 1$ which shows that we have a duplication of any subset of size s less than t. Therefore, for every concept C in $\mathcal{C} = \mathcal{C}_{t\text{-}(v,k,1)}$ one can consider any t-subset as a smallest teaching set, i.e., $\mathrm{TD}^+(C, \mathcal{C}) = t$ which implies $\mathrm{TD}^+(\mathcal{C}) = \mathrm{RTD}^+(\mathcal{C}) = t$ and consequently $\mathrm{NCTD}^+(\mathcal{C}) \le t$. □

Finally, let us turn to the objective of constructing a family of concept classes for which the gap between RTD and NCTD is arbitrarily large. With our notation, let $\mathcal{C}_{H,n}$ be a concept class built based on the Hadamard matrices H_n of order 2^n, $n \ge 1$, with zeroes replacing the -1 entries such that the rows are concepts and the columns are instances. Therefore, we have:

$$\mathcal{C}_{H,n+1} = \begin{bmatrix} \mathcal{C}_{H,n} & \mathcal{C}_{H,n} \\ \mathcal{C}_{H,n} & \mathcal{C}_{\widehat{H},n} \end{bmatrix},$$

where $\mathcal{C}_{\widehat{H},n}$ results from $\mathcal{C}_{H,n}$ by flipping every 1 to a 0 and vice versa.

Since the VCD is an essential learning complexity parameter, one goal here is also determining the VCD for the particular design $\mathcal{C}_{H,n}$. In [35], it was shown that for any t-(v, k, λ) design the relation $t \le \mathrm{VCD} \le k$ holds. As a consequence we have the following result.

Proposition 3.2.15 *For all $n \geq 2$,* $\text{VCD}(\mathcal{C}_{H,n}) = n$*.*

Proof: By induction on n. For $n = 1$ and $n = 2$ the statement is obviously true. Now let $\text{VCD}(\mathcal{C}_{H,n}) = n$ hold for a fixed value of n. Since $\mathcal{C}_{H,n}$ is a subclass of $\mathcal{C}_{H,n+1}$ and also $\text{VCD}(\mathcal{C}_{H,n})$ $=\text{VCD}(\mathcal{C}_{\widehat{H},n}) = n$, we conclude that $\text{VCD}(\mathcal{C}_{H,n+1}) \geq n$, which means there is a subset $\mathcal{X}'$ of the left half of the domain of $\mathcal{C}_{H,n+1}$, say $\mathcal{X}$, associated to the concept class $(\mathcal{C}_{H,n})$ of size n such that $|\{C|_{\mathcal{X}'} : C \in \mathcal{C}_{H,n+1}\}| = 2^n$. Now consider the subset $\mathcal{X}'' = \mathcal{X}' \cup \{x\}$, where x is the instance of $\mathcal{X}$ corresponding to the first column of the right half of $\mathcal{C}_{H,n+1}$. Hence $|\{C|_{\mathcal{X}''} : C \in \mathcal{C}_{H,n+1}\}| = 2^{|\mathcal{X}''|} = 2^{n+1}$. On the other hand, since $|\mathcal{C}_{H,n+1}| = 2^{n+1}$, $n+1$ is the maximum size of a shattered subset of $\mathcal{X}$, and therefore $\text{VCD}(\mathcal{C}_{H,n+1}) = n+1$. □

According to the previous observations, it follows that $\text{NCTD}(\mathcal{C}_{H,3}) = 1$, while $\text{RTD}(\mathcal{C}_{H,3}) = 3$, which shows again that RTD can exceed NCTD by a factor of 3. The same construction provides our first and so far only example in which RTD exceeds NCTD by a factor of 4, namely $\text{NCTD}(\mathcal{C}_{H,4}) = 1$, while $\text{RTD}(\mathcal{C}_{H,4}) = 4$. When increasing n further, we obtain $\text{NCTD}(\mathcal{C}_{H,5}) = 2$, while $\text{RTD}(\mathcal{C}_{H,5}) = 5$, which decreases the ratio of RTD over NCTD again. However, it is conceivable that the sequence of concept classes $\mathcal{C}_{H,n}$ might provide an arbitrarily large ratio between RTD and NCTD. Since the computational complexity of computing the NCTD of a given concept class is too high for us to check concept classes of bigger sizes, the general relationship between n and $\text{NCTD}(\mathcal{C}_{H,n})$ is still unknown and worthwhile to consider as an open problem. In addition, we conjecture that $\mathrm{RTD}(\mathcal{C}_{H,n}) = \mathrm{NCTD}^{+}(\mathcal{C}_{H,n}) = n$.

3.3 Enumerating All Concept Classes Over at Most Five Instances

For computational support, this section describes the implementation of a python toolbox to compute teaching complexity parameters. The toolbox enables us to analyze concept classes whose structure is more sophisticated to investigate and to compute such parameters for a large number of concept classes.

Recall that the main question of this chapter was finding concept classes $\mathcal{C}$ for which the ratio of RTD over NCTD is greater than 2. To find more such concept classes, we first wrote a python program, based on the provided algorithm in Appendix 6.3, to compute the recursive teaching dimension of a concept class $\mathcal{C}$, and then, with the help of existing code for computing the non-clash teaching dimension [36], we succeeded to generate numerous concept classes for which our well-known ratio is at least 3. We wanted this list of concept classes to be duplicate-free in the sense that no two classes in the list are equivalent. Two concept classes are equivalent if their matrices can be obtained from each other by permuting rows, permuting columns, and flipping all labels in any subset of columns [40].

According to Theorem 3.1.2, the concept class $\mathcal{C}_{3/1}$, illustrated in Table 3.2, is the smallest concept class with RTD = 3 and NCTD = 1. Therefore, it is reasonable to check all non-equivalent concept classes over 4 instances in the hope of finding another concept class that differs from $\mathcal{C}_{3/1}$ or equivalently from $\mathcal{C}'_{3/1}$.

Generally, the corresponding algorithm is able to take a list of concept classes of size n over the domain with m instances and find all non-equivalent classes with a running time of $O\left(2^m \cdot n!\right)$ which is not optimal. However, this algorithm suffices for our purposes here.

We found 74 non-equivalent concept classes of size 8 over 4 instances, including $\mathcal{C}'_{3/1}$ as well, listed in Appendix 6. Out of these distinct 8 by 4 concept classes, $\mathcal{C}'_{3/1}$ was the only concept class with the desired teaching complexity values, i.e. $\mathcal{C}'_{3/1}$ is unique. Over 4 instances, any other concept class $\mathcal{C}$ with more than 8 concepts has NCTD = 2 with RTD at most 4 so that their ratio is at most 2.

Over 5 instances, concept classes of size 8, 9 and 10 have a potential to display our desired ratios.

Regarding the 8 by 5 case, we could randomly generate only three non-equivalent concept classes, see Table 3.7, which show that they all extended the concept class $\mathcal{C}'_{3/1}$ by one additional column, as indicated by the black box. In fact, adding an extra instance x_5 to the domain of $\mathcal{C}'_{3/1}$ does not affect any underlying teaching sets associated to the concepts in $\mathcal{C}'_{3/1}$.

Similar to the 8 by 5 case, there are only three non-equivalent concept classes of size 9

	x_1	x_2	x_3	x_4	x_5
C_1	1	1	1	1	1
C_2	1	0	1	0	1
C_3	1	1	0	0	1
C_4	1	0	0	1	1
C_5	0	0	0	0	1
C_6	0	1	0	1	1
C_7	0	0	1	1	1
C_8	0	1	1	0	1

$\mathcal{C}_1^{8\times5}$

	x_1	x_2	x_3	x_4	x_5
C_1	1	1	1	1	0
C_2	1	0	1	0	0
C_3	1	1	0	0	0
C_4	1	0	0	1	0
C_5	0	0	0	0	1
C_6	0	1	0	1	1
C_7	0	0	1	1	1
C_8	0	1	1	0	1

$\mathcal{C}_2^{8\times5}$

	x_1	x_2	x_3	x_4	x_5
C_1	1	1	1	1	0
C_2	1	0	1	0	1
C_3	1	1	0	0	0
C_4	1	0	0	1	1
C_5	0	0	0	0	0
C_6	0	1	0	1	1
C_7	0	0	1	1	0
C_8	0	1	1	0	1

$\mathcal{C}_3^{8\times5}$

Table 3.7: Three non-equivalent concept classes of size 8 over 5 instances with RTD = 3 and NCTD = 1.

	x_1	x_2	x_3	x_4	x_5
C_1	1	1	1	1	1
C_2	1	0	1	0	1
C_3	1	1	0	0	1
C_4	1	0	0	1	1
C_5	0	0	0	0	1
C_6	0	1	0	1	1
C_7	0	0	1	1	1
C_8	0	1	1	0	1
C_5'	0	0	0	0	0

$\mathcal{C}_1^{9\times5}$

	x_1	x_2	x_3	x_4	x_5
C_1	1	1	1	1	1
C_2	1	0	1	0	1
C_3	1	1	0	0	1
C_4	1	0	0	1	1
C_5	0	0	0	0	1
C_6	0	1	0	1	1
C_7	0	0	1	1	1
C_8	0	1	1	0	1
C_9	1	0	0	0	0

$\mathcal{C}_2^{9\times5}$

	x_1	x_2	x_3	x_4	x_5
C_1	1	1	1	1	0
C_2	1	0	1	0	0
C_3	1	1	0	0	0
C_4	1	0	0	1	0
C_5	0	0	0	0	1
C_6	0	1	0	1	1
C_7	0	0	1	1	0
C_8	0	1	1	0	1
C_7'	0	0	1	1	1

$\mathcal{C}_3^{9\times5}$

Table 3.8: Three non-equivalent concept classes of size 9 over 5 instances with RTD = 3 and NCTD = 1.

over the same domain, see Table 3.8. These concept classes have been taken from the list of all 9,013 distinct non-equivalent concept classes produced by [29]. In addition, out of all 19,963 distinct non-equivalent concept classes of size 10 over 5 instances, there is only one concept class, shown in Table 3.9, for which the non-clash teaching model requires only teaching sets of size one, while the recursive teaching dimension is 3.

	x_1	x_2	x_3	x_4	x_5
C_1	1	1	1	1	0
C_2	1	0	1	0	0
C_3	1	1	0	0	0
C_4	1	0	0	1	1
C_5	0	0	0	0	1
C_6	0	1	0	1	1
C_7	0	0	1	1	1
C_8	0	1	1	0	0
C'_3	1	1	0	0	1
C'_7	0	0	1	1	0

Table 3.9: The unique concept class $\mathcal{C}^{10\times 5}$ of size 10 over 5 instances with RTD = 3 and NCTD = 1.

	x_1	x_2	x_3	x_4	x_5	x_6
C_1	**1**	1	1	1	1	1
C_2	0	**1**	0	1	0	1
C_3	0	1	**1**	0	0	0
C_4	0	0	1	**1**	1	1
C_5	0	1	1	0	**1**	0
C_6	0	0	0	0	0	**1**
C_7	**0**	0	0	0	0	0
C_8	1	**0**	1	0	1	0
C_9	1	0	**0**	1	1	1
C_{10}	1	1	0	**0**	0	0
C_{11}	1	0	0	1	**0**	0
C_{12}	1	1	0	1	1	**0**

Table 3.10: A concept class $\mathcal{C}^{12\times 6}$ of size 12 over 6 instances with RTD = 3 and NCTD = 1.

Over 5 instances, any concept class $\mathcal{C}$ with the number of concepts ranging from 11 to 16 has NCTD = 2 with RTD at most 4, and for the remaining concept classes NCTD = 3 and RTD $\leq$ 5. Therefore, the ratio is at most 2 for all of them.

The structure of the concept classes in Tables 3.7, 3.8 and 3.9, provides some hints on how to generate a few non-equivalent concept classes of size 10 and 11 over

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
C_1	**1**	1	1	1	1	1	1
C_2	0	**1**	0	1	1	0	0
C_3	0	1	**1**	0	0	0	1
C_4	0	0	1	**1**	0	0	0
C_5	0	0	1	1	**1**	1	0
C_6	0	1	1	1	0	**1**	0
C_7	0	1	0	1	1	1	**1**
C_8	**0**	0	0	0	0	0	0
C_9	1	**0**	1	0	0	1	1
C_{10}	1	0	**0**	1	1	1	0
C_{11}	1	1	0	**0**	1	1	1
C_{12}	1	1	0	0	**0**	0	1
C_{13}	1	0	0	0	1	**0**	1
C_{14}	1	0	1	0	0	0	**0**

Table 3.11: A concept class $\mathcal{C}^{14\times 7}$ of size 14 over 7 instances with RTD = 3 and NCTD = 1.

6 instances. In addition, according to the uniqueness of the concept classes $\mathcal{C}_{3/1}$ and $\mathcal{C}^{10\times 5}$, we conjecture that if for a concept class $\mathcal{C}$ over the domain $\mathcal{X}$ the equality $|\mathcal{C}| = 2|\mathcal{X}|$ holds and $\frac{\text{RTD}(\mathcal{C})}{\text{NCTD}(\mathcal{C})} > 2$, then $\mathcal{C}$ is unique up to equivalence.

We have generated three concept classes over the domains of sizes 6, 7 and 8, shown in Tables 3.10, 3.11 and 3.12, respectively. We observe that one can divide each of them into two subclasses which are complements of each other.

It is worth noting that both concept classes $\mathcal{C}'_3$ and $\mathcal{C}^{16\times 8}$, provided in Table 3.12 and Observation 3.2, respectively, contain 16 concepts over the domain of size 8 with NCTD = 1, but, RTD($\mathcal{C}'_3$) = 4 and RTD($\mathcal{C}^{16\times 8}$) = 3. Thus, they are non-equivalent. Additionally, another two non-equivalent concept classes can be derived by considering the concept class $\mathcal{C}^{16\times 16}$ in Table 3.13, and the concept class $\mathcal{C}_{2\text{-}(16,6,2)}$ mentioned in page 36 whose RTD and NCTD are 3 and 1, respectively.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
C_1	**1**	1	1	1	1	1	1	1
C_2	0	**1**	1	1	0	0	1	1
C_3	0	0	**1**	1	1	1	0	0
C_4	1	0	0	**1**	1	0	0	0
C_5	0	1	0	0	**1**	0	0	0
C_6	0	1	0	1	1	**1**	1	0
C_7	1	0	1	1	1	0	**1**	1
C_8	0	0	1	1	1	1	0	**1**
C_9	**0**	0	0	0	0	0	0	0
C_{10}	1	**0**	0	0	1	1	0	0
C_{11}	1	1	**0**	0	0	0	1	1
C_{12}	1	1	1	**0**	0	1	1	1
C_{13}	1	0	1	1	**0**	1	1	1
C_{14}	1	0	1	0	0	**0**	0	1
C_{15}	0	1	0	0	0	1	**0**	0
C_{16}	0	1	0	0	0	0	1	**0**

Table 3.12: A concept class $\mathcal{C}^{16\times 8}$ of size 16 over 8 instances with RTD $= 3$ and NCTD $= 1$.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}
C_1	0	**1**	1	1	1	1	0	0	0	0	0	0	0	0	0	0
C_2	1	0	0	0	0	0	**1**	1	1	1	0	0	0	0	0	0
C_3	1	0	0	0	0	0	1	0	0	0	**1**	1	1	0	0	0
C_4	1	0	0	0	0	0	0	**1**	0	0	1	0	0	1	1	0
C_5	**1**	0	0	0	0	0	0	0	1	0	0	1	0	1	0	1
C_6	1	0	0	0	0	0	0	0	0	**1**	0	0	1	0	1	1
C_7	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	**1**
C_8	0	1	0	1	0	0	0	0	0	0	0	1	**1**	0	0	1
C_9	0	1	0	0	1	0	0	0	0	0	1	0	1	0	**1**	0
C_{10}	0	1	0	0	0	1	0	0	0	0	1	1	0	**1**	0	0
C_{11}	0	0	1	**1**	0	0	0	0	1	1	0	0	0	0	0	1
C_{12}	0	0	**1**	0	1	0	0	1	0	1	0	0	0	0	1	0
C_{13}	0	0	1	0	0	1	0	1	**1**	0	0	0	0	1	0	0
C_{14}	0	0	0	1	**1**	0	1	0	0	1	0	0	1	0	0	0
C_{15}	0	0	0	1	0	1	1	0	1	0	0	**1**	0	0	0	0
C_{16}	0	0	0	0	1	**1**	1	1	0	0	1	0	0	0	0	0

Table 3.13: A concept class $\mathcal{C}^{16\times 16}$ of size 16 over 16 instances with RTD $= 3$ and NCTD $= 1$.

During our research, a few other 2-designs have also been considered whose RTD and NCTD are 3 and 1, respectively. They are listed in Appendix 6.2.

Overall, in this chapter, we affirmatively answered the question whether there exists a concept class for which $\frac{\text{RTD}}{\text{NCTD}} > 2$ holds. While most of the non-equivalent concept classes obtained in this chapter have RTD = 3 and NCTD = 1, so far, there is only one concept class, $\mathcal{C}_{H,4}$, for which RTD = 4 and NCTD = 1. In addition, there is only one 32 by 16 concept class, $\mathcal{C}_{H,5}$ for which RTD = 5 and NCTD = 2. We conjecture that further investigations on Hadamard matrices may help to increase the gap between RTD and NCTD, and we leave it as an open problem.

Chapter 4

A Graph-theoretic View of Teaching

Studying the structure of specific graphs is often helpful in obtaining more general facts about graphs. Some such facts have been found by investigating matrices assigned to a graph [6]. Given that we have already made interesting connections between teaching and Hadamard matrices, it seems worthwhile to conduct an investigation in order to reveal properties of a concept class by looking at graph-theoretic notions related to concept classes.

4.1 Encoding Teaching Information in a Graph

Introduced formally in [3], the *one-inclusion graph* of a concept class is a helpful tool when it comes to studying some notions in learning theory [11, 18–20]. Research works such as [9, 25] have also established some links between the one-inclusion graph of certain concept classes and their corresponding teaching complexity including VC-dimension, classical and recursive teaching dimensions.

The main result of this chapter features a connection between the one-inclusion graph of concept classes and the non-clash teaching complexity. We found that for so-called shortest-path-closed concept classes with NCTD equal to 1, the corresponding one-inclusion graphs have at most one cycle, and consequently the recursive teaching dimension of such a class is at most 2.

In the following, we give the definition of this particular graph.

Definition 4.1.1 *The one-inclusion graph of a concept class $\mathcal{C}$, is the graph $G(C)$*[1] *for which the vertex set V is equal to $\mathcal{C}$ and the edge set E equals $\{CC' : C, C' \in \mathcal{C} \text{ and } |C \triangle C'| = 1\}$. Every edge CC' is labelled by the instance in the set $C \triangle C'$.*

As an example, consider the concept class $\mathcal{C}$ used for illustration in [9] and shown in Fig. 4.1. In the one-inclusion graph, there is an edge between two concepts as vertices if they disagree exactly on one instance. For example, concepts C_1 and C_5 differ only on instance x_2, thus in the corresponding one-inclusion graph, they are connected by an edge labelled with x_2. By repeatedly applying this rule, we obtain the graph G as shown in Fig. 4.1.

The set of one-inclusion incident instances of the concept C in $\mathcal{C}$, denoted by $I_G(C)$, is formed by collecting all instances that appear on edges incident to vertex

[1] In the following, we use G to simplify the notation $G(C)$ if unambiguous.

	x_1	x_2	x_3	x_4
C_1	0	0	0	1
C_2	0	0	1	0
C_3	0	0	1	1
C_4	0	1	0	0
C_5	0	1	0	1
C_6	0	1	1	0
C_7	0	1	1	1
C_8	1	0	0	1
C_9	1	0	1	0
C_{10}	1	1	0	0

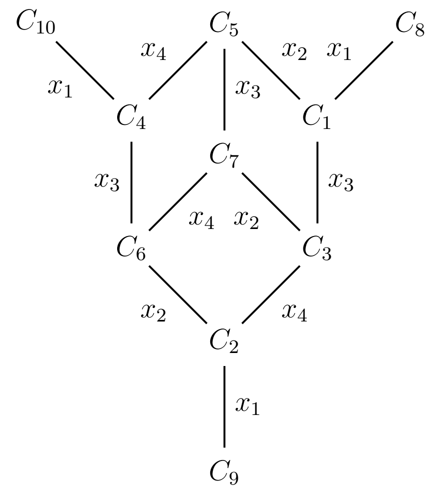


Figure 4.1: A concept class $\mathcal{C}$ and its corresponding one-inclusion graph.

C in the one-inclusion graph G, [20, 25]. Applying Proposition 3.1 in [9], which states that the size of $I_G(C)$ and the degree of the corresponding vertex of C in G are identical, shown in [9, Lemma 3.2], $\deg_G(C)$ is a lower bound on the teaching dimension of C in $\mathcal{C}$. In other words, a teaching set must at least distinguish the concept from its adjacent concepts. In some cases, it can be seen easily that the sample set corresponding to a concept C obtained by only considering its incident instances in the one-inclusion graph G is a subset of the teaching set, i.e. we may need to add more information to this sample set in order to make it unique to distinguish C from other concepts in $\mathcal{C}$. For clarity, let's consider the concept class $\mathcal{C}$ in Fig. 4.1 again. We have $T(C_1) = \{(x_1, 0), (x_2, 0), (x_3, 0)\}$ and $I_G(C_1) = \{x_1, x_2, x_3\}$. Therefore, $I_G(C_1)$ completely provides the corresponding teaching set for C_1. On the other hand, for C_{10}, $T(C_{10}) = \{(x_1, 1), (x_2, 1)\}$ and $I_G(C_{10}) \subset \{x_1, x_2\}$ which means having only the sample set $\{(x_1, 1)\}$ would not be enough to distinguish the concept C_{10} from all other concepts. These incompatibilities occur once the underlying one-inclusion graph does not have a property called *Hamming-connectivity*, which was originally defined in [9], and is stated here for completeness.

Definition 4.1.2 *Consider the concept class $\mathcal{C}$ and let $\mathcal{G} = (V, E)$ be a graph for which $V = \mathcal{C}$. Two concepts C and $C' \in V$ are Hamming-connected in $\mathcal{G}$ or $\mathcal{G}$-Hamming-connected if and only if $d_{\mathcal{G}}(C, C') = |C \triangle C'|$. The shortest path between C and C' in $\mathcal{G}$ is called a Hamming path between C and C'. In addition, a concept C is $\mathcal{G}$-fully-Hamming-connected if and only if it is $\mathcal{G}$-Hamming-connected to every $C' \in \mathcal{C}$.*

The following result, referred to as Theorem 3.8 in [9], demonstrates that for a fully G-Hamming-connected concept, the one-inclusion vertex degree and the corresponding teaching dimension are equal. Moreover, the concept's minimum teaching set will be uniquely determined.

Theorem 4.1.3 *Let $\mathcal{C}$ be a concept class and $C \in \mathcal{C}$. If C is fully G-Hamming-connected, then*

1. $\mathrm{TD}(C, \mathcal{C}) = \deg_G(C)$, *and*

2. *C has a unique minimum teaching set $T(C)$ which fulfills $\mathcal{X}(T(C)) = I_G(C)$.*

Regarding the one-inclusion graph G in Fig. 4.1, as we discussed above, since concept C_1 is fully G-Hamming-connected, its teaching dimension is exactly 3 and it has a unique teaching set formed by the set $I_G(C_1)$ of incident instances. On the other hand, as C_{10} is not G-Hamming-connected, it is not possible to form a teaching set only by its one-inclusion incident instance x_1 and consequently we need more instances to compute its teaching dimension which in fact is 2.

Theorem 4.1.3 shows how a fully G-Hamming-connected concept is helpful in recognizing the corresponding teaching set. Therefore, it is worthwhile to go beyond Definition 4.1.2 and investigate a concept class in which all concepts are fully G-Hamming-connected. Referring to the articles [25, 34], such a concept class is called *shortest-path-closed*. We state the definition here:

Definition 4.1.4 *Let $\mathcal{C}$ be a concept class and $\mathcal{G} = (V, E)$ be a graph for which $V = \mathcal{C}$. $\mathcal{C}$ is fully Hamming-connected in $\mathcal{G}$ or fully $\mathcal{G}$-Hamming-connected if and only if every concept $C \in \mathcal{C}$ is fully $\mathcal{G}$-Hamming-connected.*

Definition 4.1.5 *A concept class $\mathcal{C}$ is shortest-path-closed if and only if it is $\mathcal{G}$-fully Hamming-connected.*

While an eager reader can find a variety of interesting properties related to shortest-path-closed classes in [9], the following result determines a necessary and sufficient condition for a class to be shortest-path-closed.

Theorem 4.1.6 *Let $\mathcal{C}$ be a concept class. Then $\mathcal{C}$ is shortest-path-closed if and only if every concept $C \in \mathcal{C}$ has a unique minimum teaching set $T(C)$, where $\mathcal{X}(T(C)) = I_G(C)$.*

4.2 On Shortest-Path-Closed Classes With NCTD $= 1$

Given the characterization of shortest-path-closed classes in terms of teaching sets (Theorem 4.1.6), we now turn to studying connections between the NCTD and the one-inclusion graph. The next theorem categorizes all shortest-path-closed classes with non-clash teaching dimension 1.

Theorem 4.2.1 *Let $\mathcal{C}$ be shortest-path-closed. Then the following statements are equivalent:*

- $\mathrm{NCTD}(\mathcal{C}) = 1$*;*
- *The one-inclusion graph G corresponding to $\mathcal{C}$ has at most one cycle.*

Proof Using the handshaking property for the one-inclusion graph $G = (\mathcal{C}, E)$ and the inequality $\mathrm{NCTD}(\mathcal{C}) \geq \lceil \frac{1}{2} \cdot \deg_{avg}(\mathcal{C}) \rceil$, we have $\mathrm{NCTD}(\mathcal{C}) \geq \lceil \frac{|E|}{|\mathcal{C}|} \rceil$. $\mathrm{NCTD}(\mathcal{C}) = 1$ results in the inequality $|E| \leq |\mathcal{C}|$. On the other hand, since the one-inclusion graph G is connected, we conclude that $|E| \geq |\mathcal{C}| - 1$. Therefore, we have either $|E| = |\mathcal{C}| - 1$ or $|E| = |\mathcal{C}|$. If $|E| = |\mathcal{C}| - 1$, then the one-inclusion graph must be a tree. If $|E| = |\mathcal{C}|$, one can see that the one-inclusion graph contains exactly one cycle.

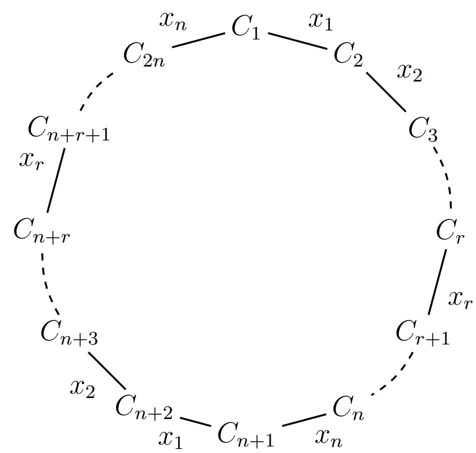


Figure 4.2: The one-inclusion graph G corresponding to the concept class $\mathcal{C}_{\mathfrak{C}_{2n}}$.

Conversely, if G is tree, then it is easy to see that $\mathrm{RTD}(\mathcal{C}) = 1$ [7] and consequently $\mathrm{NCTD}(\mathcal{C}) = 1$. Suppose that G has only one cycle, named $\mathfrak{C}_k$, of size k. Since this cycle is a closed path, each involved instance needs to occur an even number of times as an edge label. Moreover, as $\mathcal{C}$ is shortest-path-closed, this even number must always be 2, which also proves that k is an even number. Let $k = 2n$, $n \geq 2$, and suppose $\mathcal{X}_{\mathfrak{C}_{2n}} = \{x_1, x_2, \ldots, x_n\}$ is the set of instances labelling edges in $\mathfrak{C}_{2n}$ with the vertex set $\mathcal{C}_{\mathfrak{C}_{2n}} = \{C_1, C_2, \ldots, C_{2n}\}$, where $\mathfrak{C}_{2n} = (C_1, C_2, \ldots, C_{2n}, C_1)$.

Now to make the two concepts C_1 and C_{n+1} Hamming-connected in G, each instance should appear only once in a shortest path between them. However, this path is of maximum size n and thus needs to contain instances x_1, x_2, $\ldots$, x_n as edge labels. Moving forward one by one, by considering all pairs (C_r, C_{n+r}) and the shortest paths between them, shows that the set of edge labels on any such shortest path must always be $\{x_1, \ldots, x_n\}$, and therefore, we conclude that the induced sub-one-inclusion graph corresponding to the concept class $\mathcal{C}_{\mathfrak{C}_{2n}}$, illustrated in Fig. 4.2, is unique. Without loss of generality, suppose that all labels for the concept C_1 are 0. Therefore, we have the below concept class $\mathcal{C}_{\mathfrak{C}_{2n}}$ assigned to the one-inclusion graph G in Fig. 4.2.

$$
\mathcal{C}_{\mathfrak{C}_{2n}} : \quad
\begin{array}{c}
 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ \vdots \\ C_r \\ \vdots \\ C_n \\ C_{n+1} \\ C_{n+2} \\ C_{n+3} \\ C_{n+4} \\ \vdots \\ C_{n+r} \\ \vdots \\ C_{2n}
\end{array}
\begin{array}{c}
\begin{array}{cccccccccc}
x_1 & x_2 & x_3 & x_4 & \ldots & x_{r-1} & x_r & \ldots & x_{n-1} & x_n
\end{array} \\
\left[\begin{array}{cccccccccc}
\textcircled{0} & 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\
1 & \textcircled{0} & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\
1 & 1 & \textcircled{0} & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\
1 & 1 & 1 & \textcircled{0} & \ldots & 0 & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \ldots & \ddots & \ldots & \ldots & \vdots & \vdots \\
1 & 1 & 1 & 1 & \ldots & 1 & \textcircled{0} & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ldots & \ddots & \ddots & \ddots & \ddots & \vdots \\
1 & 1 & 1 & 1 & \ldots & 1 & 1 & 1 & 1 & \textcircled{0} \\
\textcircled{1} & 1 & 1 & 1 & \ldots & 1 & 1 & 1 & 1 & 1 \\
0 & \textcircled{1} & 1 & 1 & \ldots & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & \textcircled{1} & 1 & \ldots & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & \textcircled{1} & \ldots & 1 & 1 & 1 & 1 & 1 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ldots & \ldots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \ldots & 0 & \textcircled{1} & \ldots & 1 & 1 \\
\vdots & \vdots & \vdots & \vdots & \ldots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \textcircled{1}
\end{array}\right]
\end{array},
$$

Note that the sample sets $\{(x_r, 0)\}$ and $\{(x_{n+r}, 1)\}$ form non-clash teaching sets for the concepts C_r and C_{n+r}, $1 \leq r \leq n$, respectively, with respect to $\mathcal{C}_{\mathfrak{C}_{2n}}$.

Now suppose T_j is a tree in $\mathcal{C}$ attached to the vertex C_j, $1 \leq j \leq 2n$, and disjoint from the cycle $\mathfrak{C}_{2n}$. With the same argument as before, we are only allowed to use different labels on each shortest path. Therefore, for each $1 \leq j \leq 2n$ there is a set $\mathcal{X}_{T_j} = \{x_1^j, x_2^j, \ldots, x_{|T_j|-1}^j\}$ of size $|T_j| - 1$ such that each instance in $\mathcal{X}_{T_j}$ is an edge label for exactly one edge in T_j. In addition, for all $i \neq j \in \{1, 2, \ldots, 2n\}$, $\mathcal{X}_{\mathcal{C}_{2n}} \cap \mathcal{X}_{T_i} \cap \mathcal{X}_{T_j} = \emptyset$ holds.

In case each vertex on the cycle $\mathfrak{C}_{2n}$ has exactly one such tree attached to it, the concept class $\mathcal{C}$ has the following block structure:

$$
\mathcal{C} : \quad
\begin{array}{c}
\\ \\
\end{array}
\begin{array}{c|ccccc}
 & \mathcal{X}_{\mathfrak{C}_{2n}} & \mathcal{X}_{T_1} & \ldots & \mathcal{X}_{T_j} & \ldots & \mathcal{X}_{T_{2n}} \\
\hline
 & \boxed{\mathfrak{C}_{2n}} & 0_1 & \ldots & 0_j & \ldots & 0_{2n} \\
 & \mathcal{C}_{C_1} & \boxed{\mathcal{C}_{T_1}} & & & & \\
 & \vdots & & \ddots & & 0_2 & \\
 & \mathcal{C}_{C_j} & & & \boxed{\mathcal{C}_{T_j}} & & \\
 & \vdots & & 0_3 & & \ddots & \\
 & \mathcal{C}_{C_{2n}} & & & & & \boxed{\mathcal{C}_{T_{2n}}}
\end{array}
\quad \rightarrow M
$$

where

$$
\mathcal{C}_{T_j} : \quad
\begin{array}{c}
\\
C_1^j \\ C_2^j \\ C_3^j \\ \vdots \\ C_s^j \\ \vdots \\ C_{|T_j|-1}^j
\end{array}
\begin{array}{c}
\begin{array}{ccccccc}
x_1^j & x_2^j & x_3^j & \ldots & x_s^j & \ldots & x_{|T_j|-1}^j
\end{array} \\
\left[
\begin{array}{ccccccc}
\textcircled{1} & & & & & & \\
 & \textcircled{1} & & & & & \\
 & & \textcircled{1} & & & 0 & \\
 & & & \ddots & & & \\
 & & & & \textcircled{1} & & \\
 & & * & & & \ddots & \\
 & & & & & & \textcircled{1}
\end{array}
\right]_{|T_j|-1 \times |T_j|-1}
\end{array}
,
$$

$$\mathcal{C}_{C_j} : \begin{array}{c} \begin{matrix} x_1 & x_2 & \dots & x_{j-1} & x_j & \dots & x_n \end{matrix} \\ \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & \dots & 1 & 0 & \dots & 0 \end{bmatrix}_{|T_j|-1 \times n} \end{array} \quad \text{and } \mathcal{C}_{C_{n+j}} = \hat{\mathcal{C}}_{C_j}.$$

Non-clash teaching sets for the concepts C_s^j are $\{(x_s^j, 1)\}$, $1 \leq s \leq |T_j| - 1$ and $1 \leq j \leq 2n$. In addition, teaching sets for $\mathcal{C}_{\mathfrak{C}_{2n}}$ and $\mathcal{C}_{T_j}$, $1 \leq j \leq 2n$, form non-clash teaching sets for the concept class $\mathcal{C}$. For (subclass) sub-matrix M, both zero-parts 0_2 and 0_3 help to show that there is no clash between teaching sets $\{(x_s^j, 1)\}$ corresponding to the $\mathcal{C}_{T_j}$'s. Sub-matrices 0_j of order $2n \times |T_j| - 1$ prove that the teaching sets for $\mathcal{C}_{\mathfrak{C}_{2n}}$ have no clash with $\mathcal{C}_{T_j}$'s teaching sets. Therefore, $\mathrm{NCTD}(\mathcal{C}) = 1$. In case a vertex C_j on the cycle has no tree attached to it, one simply removes the corresponding block in the above matrix. Likewise, if C_j has multiple trees attached to it, the same argument goes through when using more than one block corresponding to C_j. □

Corollary 4.2.2 *Let $\mathcal{C}$ be a shortest-path-closed concept class. If* $\mathrm{NCTD}(\mathcal{C}) = 1$, *then* $\mathrm{RTD}(\mathcal{C})$ *is at most 2.*

Here is an example to illustrate the proof of Theorem 4.2.1.

Example 4.2.3 *Consider the one-inclusion graph G that includes vertices C_1, C_2, ..., C_{14} and the edge label set $\{x_1, \dots, x_{12}\}$, as depicted in Fig. 4.3.*

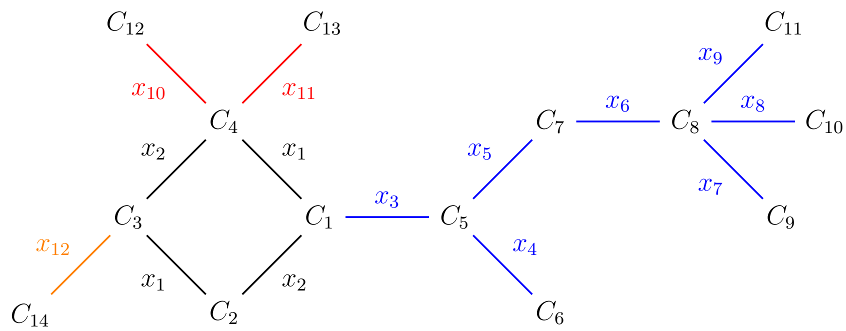


Figure 4.3: A one-inclusion graph G which has only one cycle of size 4.

Below, we list the underlying concept class $\mathcal{C}$ corresponding to the above graph, where the non-clash teaching sets are all circled.

	$\mathcal{X}_{\mathfrak{C}_4}$		$\mathcal{X}_{T_1}$							$\mathcal{X}_{T_4}$		$\mathcal{X}_{T_3}$
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
C_1	0	0	0	0	0	0	0	0	0	0	0	0
C_2	1	0	0	0	0	0	0	0	0	0	0	0
C_3	1	1	0	0	0	0	0	0	0	0	0	0
C_4	0	1	0	0	0	0	0	0	0	0	0	0
C_5	0	0	1	0	0	0	0	0	0	0	0	0
C_6	0	0	1	1	0	0	0	0	0	0	0	0
C_7	0	0	1	0	1	0	0	0	0	0	0	0
C_8	0	0	1	0	1	1	0	0	0	0	0	0
C_9	0	0	1	0	1	1	1	0	0	0	0	0
C_{10}	0	0	1	0	1	1	0	1	0	0	0	0
C_{11}	0	0	1	0	1	1	0	0	1	0	0	0
C_{12}	0	1	0	0	0	0	0	0	0	1	0	0
C_{13}	0	1	0	0	0	0	0	0	0	0	1	0
C_{14}	1	1	0	0	0	0	0	0	0	0	0	1

M

In this chapter, the notion of one-inclusion graph was investigated. Moreover, some definitions such as G-Hamming-connectivity and shortest-path-closed classes were introduced. The one-inclusion graph structure is based on labels of size one, i.e.

the concepts that disagree on exactly one instance are connected by an edge. Therefore, a question that arises here is: What can one say if two concepts have different labels on multiple instances? In [9], it was shown that the Hamming-connectivity of a graph in these cases is also a strong tool to establish a link between the classical teaching model and the graph representation, and consequently, reveals information about the teaching complexity of a concept class. Therefore, studies akin to [9] can lead to new results related to the non-clash teaching dimension.

Chapter 5

Conclusions and Further Work

To summarize, Chapter 1 of this thesis introduced some of the best-known research works regarding certain teaching settings such as the batch teaching models and the sequential teaching models. Such well-studied research works were our motivation to investigate those concept classes that are easy to teach in the strongest batch model, i.e., non-clashing model [24], more in depth. Although the non-clashing model is optimal among all models satisfying Goldman and Mathias's notion of collusion-freeness, we were previously only aware of concept classes for which the ratio of RTD and NCTD is at most 2.

In Section 3.1, we built step by step a novel concept class of size 8 over a domain includes 8 instances for which the ratio of RTD and NCTD is at most 3. We could then reduce the number of instances to 4 and verify the existence of an 8 by 4 concept class which fulfills RTD = 3 and NCTD = 1. This class is the smallest concept class with this property.

Then in Section 3.2, we used design theory as a tool to help increase the gap between RTD and NCTD. We found that there exists a concept class of size 16 over 8 instances that satisfies RTD = 4 and NCTD = 1. At the time of writing this thesis, this example represents the largest multiplicative gap between RTD and NCTD.

Section 3.3 describes the implementation of a python toolbox to compute teaching complexity parameters. We have obtained that over 4 instances, there is a unique concept class of size 8 which fulfills RTD = 3 and NCTD = 1. On the other hand, over 5 instances, there are three non-equivalent concept classes of size 8, and also of size 9, as well as one unique concept class of size 10 for which RTD = 3 and NCTD = 1 hold. We generated concept classes of size 12, 14 and 16 over domains of size 6, 7 and 8, respectively, for which RTD = $3 \cdot$ NCTD.

Finally, the main result of Chapter 4 states that for a given shortest-path-closed concept class $\mathcal{C}$ with NCTD equal to 1, the associated one-inclusion graph has at most one cycle, and consequently the recursive teaching dimension of $\mathcal{C}$ is at most 2.

We close the chapter by listing a few interesting items extracted from this thesis that are worthwhile to investigate in the future.

1. It can be verified that the two parameters RTD and VCD are equal for all discovered new concept classes in Chapter 3. The question that arises here is whether the equality RTD = VCD is a necessary condition for the inequality

$\mathrm{RTD}(\mathcal{C}) \geq 3 \cdot \mathrm{NCTD}(\mathcal{C})$, for any concept class $\mathcal{C}$.

2. The general relationship between n and $\mathrm{NCTD}(\mathcal{C}_{H,n})$ remains a notable open problem. By increasing n, we believe that RTD and NCTD of the concept class $\mathcal{C}_{H,n}$ grow linearly and logarithmically, respectively, which implies that an arbitrarily large ratio between RTD and NCTD could be attained. We also conjecture that $\mathrm{RTD}(\mathcal{C}_{H,n}) = \mathrm{NCTD}^+(\mathcal{C}_{H,n}) = n$.

3. Definition 3.2.13 outlines the Kronecker product notion. In Section 3.2, this operation was used to construct a concept class which fulfills the equality $\mathrm{RTD} = 4 \cdot \mathrm{NCTD}$. Therefore, generating new concept classes by applying such operations is helpful to obtain some additional results about teaching parameters. Doliwa et al. [7] used the disjoint union of two concept classes $\mathcal{C}_1$ and $\mathcal{C}_2$ to show that the difference between RTD and VCD can become arbitrarily large.

Definition 5.0.1 *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be concept classes over disjoint domains $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. Then over the domain $\mathcal{X}_1 \cup \mathcal{X}_2$, a new concept class is defined as $\mathcal{C}_1 \sqcup \mathcal{C}_2 = \{C_1 \cup C_2 : C_1 \in \mathcal{C}_1$ and $C_2 \in \mathcal{C}_2\}$.*

They proved that $\mathrm{RTD}(\mathcal{C}_1 \sqcup \mathcal{C}_2) = \mathrm{RTD}(\mathcal{C}_1) + \mathrm{RTD}(\mathcal{C}_2)$ and $\mathrm{VCD}(\mathcal{C}_1 \sqcup \mathcal{C}_2) = \mathrm{VCD}(\mathcal{C}_1) + \mathrm{VCD}(\mathcal{C}_2)$. Similarly, the inequality $\mathrm{NCTD}(\mathcal{C}_1 \sqcup \mathcal{C}_2) \leq \mathrm{NCTD}(\mathcal{C}_1) + \mathrm{NCTD}(\mathcal{C}_2)$ holds for given concept classes $\mathcal{C}_1$ and $\mathcal{C}_2$ due to Lemma 17 in [24].

Such results may motivate readers to work on the Kronecker product or another operations of two given concept classes $\mathcal{C}_1$ and $\mathcal{C}_2$ in order to establish some further results about teaching parameters.

4. In Section 3.3 the uniqueness of the concept classes $\mathcal{C}_{3/1}$ and $\mathcal{C}^{10\times 5}$ was proved. We conjecture that if for a given concept class $\mathcal{C}$ over the domain $\mathcal{X}$ the equality $|\mathcal{C}| = 2|\mathcal{X}|$ holds and $\frac{\mathrm{RTD}(\mathcal{C})}{\mathrm{NCTD}(\mathcal{C})} > 2$, then $\mathcal{C}$ is unique modulo equivalence. This remained unresolved.

5. Theorem 4.2.1 in Chapter 4 characterizes all shortest-path-closed classes with non-clash teaching dimension equal to 1. How can one classify all shortest-path-closed classes with non-clash teaching dimension of size 2 or more?

Bibliography

[1] Ethem Alpaydin. *Introduction to Machine Learning.* 4th Edition, The MIT Press, London, United Kingdom, 2020.

[2] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. Learnability and the Vapnik-Chervonenkis Dimension. *Journal of the Association for Computing Machinery*, **36**(4):929–965, 1989.

[3] John A. Bondy. Induced subsets. *Journal of Combinatorial Theory, Series B*, **12**(2):201–202, 1972.

[4] Xi Chen, Yu Cheng, and Bo Tang. On the recursive teaching dimension of VC classes. *In Proceedings of 30th Conference on Neural Information Processing Systems*, Barcelona, Spain, pp. 2164–2171, 2016.

[5] Angela Dean, Daniel Voss, and Danel Draguljic. *Design and Analysis of Experiments.* 2nd Edition, Springer International Publishing, Cham, Switzerland, 2017.

[6] Reinhard Diestel. *Graph Theory.* 5th Edition, Springer, Berlin, Germany, 2017.

[7] Thorsten Doliwa, Gaojian Fan, Hans U. Simon, and Sandra Zilles. Recursive teaching dimension, VC-dimension and sample compression. *Journal of Machine Learning Research*, **15**(1):3107–3131, 2014.

[8] Malte Darnstädt, Thorsten Kiss, Hans Ulrich Simon, and Sandra Zilles. Order compression schemes. *Theoretical Computer Science*, **620**:73–90, 2016.

[9] Gaojian Fan. *A Graph-Theoretic View of Teaching.* Master's thesis, Faculty of Graduate Studies and Research, University of Regina, Canada, 2012.

[10] Ronald A. Fisher and Frank Yates. *Statistical Tables for Biological, Agricultural and Medical Research.* Oliver and Boyd, Edinburgh, United Kingdom, 1938.

[11] Sally Floyd. *On Space-bounded Learning and the Vapnik-Chervonenkis Dimensions.* PhD thesis, International Computer Science Institute, Berkeley, CA, 1989.

[12] Sally Floyd and Manfred K. Warmuth. Sample compression, learnability, and the Vapnik-Chervonenkis dimension. *Machine Learning*, **21**(3):269–304, 1995.

[13] Martin J. Erickson. *Introduction to Combinatorics.* 2nd Edition, John Wiley & Sons Inc, New York, United States, 2013.

[14] Ziyuan Gao. *Distinguishing Linear Sets and Pattern Languages With Membership Examples*, PhD thesis, Faculty of Graduate Studies and Research, University of Regina, Canada, 2017.

[15] Ziyuan Gao, Christoph Ries, Hans U. Simon, and Sandra Zilles. Preference-based teaching. *Journal of Machine Learning Research*, **18**(31):1–32, 2017.

[16] Sally A. Goldman and Michael J. Kearns. On the complexity of teaching. *Journal of Computer and System Sciences*, **50**:20–31, 1995.

[17] Sally A. Goldman and H. David Mathias. Teaching a smarter learner. *Journal of Computer and System Sciences*, **52**(2):255–267, 1996.

[18] Leonid Gurvits. Linear algebraic proofs of VC-dimension based inequalities. *In Proceedings of the European Conference on Computational Learning Theory*, Jerusalem, Israel, pp. 238–250, 1997.

[19] David Haussler. Sphere packing numbers for subsets of the boolean n-cube with bounded Vapnik-Chervonenkis dimension. *Journal of Combinatorial Theory*, Series A, **69**(2):217–232, 1995.

[20] David Haussler, Nick Littlestone, and Manfred K. Warmuth. Predicting $\{0,1\}$-functions on randomly drawn points. *Information and Computation*, **115**(2):248–292, 1994.

[21] Lunjia Hu, Ruihan Wu, Tianhong Li, and Liwei Wang. Quadratic upper bound for recursive teaching dimension of finite VC classes. *In Proceedings of the 30th Conference on Learning Theory*, Amsterdam, The Netherlands, pp. 1147–1156, 2017.

[22] Yury J. Ionin, Mohan S. Shrikhande. *Combinatorics of Symmetric Designs.* Cambridge University Press, Cambridge, United Kingdom, 2006.

[23] Thomas P. Kirkman. On a problem in combinatorics. *Cambridge and Dublin Math Journal*, **2**:191–204, 1847.

[24] David Kirkpatrick, Hans U. Simon, and Sandra Zilles. Optimal Collusion-Free Teaching. *In Proceedings of the 30th International Conference on Algorithmic Learning Theory*, Chicago, Illinois, USA, pp. 506–528, 2019

[25] Dima Kuzmin and Manfred K. Warmuth. Unlabeled compression schemes for maximum classes. *Journal of Machine Learning Research*, **8**:2047–2081, 2007.

[26] Warwick D. Launey and Dane Flannery. *Algebraic Design Theory.* American Mathematical Society, Rhode Island, United States, 2011.

[27] Nick Littlestone and Manfred K. Warmuth. Relating data compression and learnability. *Unpublished Manuscript*, 1986.

[28] Farnam Mansouri, Yuxin Chen, Ara Vartanian, Jerry Zhu, and Adish Singla. Preference-Based Batch and Sequential Teaching: Towards a Unified View of Models. *In Proceedings of the 33rd Conference on Neural Information Processing Systems*, Vancouver, BC, Canada, pp. 9195–9205, 2019.

[29] Regan Meloche. Exploring the Relationships between Properties of Concept Classes. *Unpublished Manuscript*, University of Regina, Canada, 2014.

[30] Sudarshan R. Nagpaul and Surender K. Jain. *Topics in Applied Abstract Algebra*. Brooks/Cole, California, United States, 2005.

[31] Hans U. Simon and Sandra Zilles. Open problem: Recursive teaching dimension versus VC dimension. *In Proceedings of the 28th Conference on Learning Theory*, Paris, France, pp. 1770–1772, 2015.

[32] Ayumi Shinohara and Satoru Miyano. Teachability in computational learning. *New Generation Computing*, **8**:337–348, 1991.

[33] Douglas. R. Stinson. *Combinatorial Designs: Constructions and Analysis*. Springer-Verlag New York, New York, United States, 2004.

[34] Benjamin I. P. Rubinstein and J. Hyam Rubinstein. A geometric approach to sample compression. *Journal of Machine Learning Research*, **13**:1221–1261, 2012.

[35] Ryan Tessier. *On The Combinations of Sample Compression Schemes*. Master's thesis, Faculty of Graduate Studies and Research, University of Regina, Canada, 2014.

[36] Shayantonee D. Tupor. No Clash Teaching Model and Recursive Teaching Model. Machine learning course report. *Unpublished Manuscript*, University of Regina, Canada, 2019.

[37] Vladimir N. Vapnik and Alexey Y. Chervonenkis. On the uniform con- vergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, **16**:264–280, 1971.

[38] Leslie G. Valiant. A theory of the learnable. *Communications of the ACM*, **27**(11):1134–1142, 1984.

[39] Xiaojin Zhu, Adish Singla, Sandra Zilles, and Anna N. Rafferty. An overview of machine teaching. *https://arxiv.org/abs/1801.05927*, 2018.

[40] Sandra Zilles, Steffen Lange, Robert Holte, and Martin Zinkevich. Models of cooperative teaching and learning. *Journal of Machine Learning Research*, **12**:349–384, 2011.

Chapter 6

Appendix

6.1 74 non-equivalent concept classes of size 8 over 4 instances

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&0\\1&0&0&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&0\\0&1&1&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&0\\0&1&0&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&0\\1&1&1&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&0\\1&1&0&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&0\\0&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&0\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\0&0&1&1\\1&1&1&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\0&0&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&1&1&0\\1&1&0&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&1&1&0\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&1&1&0\\1&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&0&0\\1&0&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&1&0\\1&1&0&1\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&1&0\\1&1&0&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&1&0\\1&1&1&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&1&0\\1&1&0&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&1&0\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&0&1\\0&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&0&1\\1&1&0&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&0&1\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&0&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&0&1\\1&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&0\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&0&1\\0&1&0&1\\0&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&0&1\\0&1&0&1\\1&1&1&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&0&1\\0&1&0&1\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&0&1\\0&1&0&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&0&1\\0&0&1&1\\1&0&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&0&1\\0&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&0&1\\0&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&0&1\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\0&0&1&1\\1&1&1&0\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\0&0&1&1\\1&1&1&0\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\0&0&1&1\\1&1&0&1\\1&0&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\0&0&1&1\\1&1&0&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\0&0&1&1\\1&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\0&0&1&1\\1&0&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&1&1&0\\1&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&1&0&1\\1&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&0\\1&0&1&1\\0&1&1&1\\1&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\0&0&1&1\\1&1&1&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\0&0&1&1\\1&1&0&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\0&0&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\1&1&1&0\\1&1&0&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\1&1&1&0\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\1&1&1&0\\1&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\1&1&0&1\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\1&1&0&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\1&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\0&1&0&1\\1&0&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\1&1&1&0\\1&1&0&1\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\1&1&1&0\\1&1&0&1\\0&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\1&1&1&0\\1&1&0&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\1&1&1&0\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\1&1&0&1\\1&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\\1&1&0&1\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&1&0\\1&1&0&1\\1&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&1&0\\1&1&0&1\\1&0&1&1\\1&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&0&1\\1&0&1&1\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&1&0&0\\0&0&1&1\\1&0&1&1\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&1&0\\0&0&1&1\\1&1&1&0\\1&1&0&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&1&0\\0&0&1&1\\1&1&0&1\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&1&0\\0&0&1&1\\1&1&0&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&1&0\\0&0&1&1\\1&0&1&1\\0&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&1&0\\1&1&0&1\\1&0&1&1\\0&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&1&0\\1&1&0&1\\1&0&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&1&0\\1&0&1&1\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&0&1\\1&0&1&1\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&0&1&0\\0&0&1&1\\1&1&0&1\\0&1&1&1\\1&1&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&1&1\\1&1&1&0\\1&1&0&1\\1&0&1&1\\0&1&1&1\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&1&0\\0&1&0&1\\0&0&1&1\\1&1&1&0\\1&1&0&1\\1&0&1&1\end{bmatrix}, \begin{bmatrix}0&0&0&0\\1&1&0&0\\1&0&1&0\\1&0&0&1\\0&1&1&0\\0&1&0&1\\0&0&1&1\\1&1&1&1\end{bmatrix},$$

6.2 2-designs

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
C_1	**1**	1	1	1	1	0	0	0	0	0	0
C_2	1	1	**0**	0	0	1	1	1	0	0	0
C_3	1	0	1	**0**	0	1	0	0	1	1	0
C_4	1	0	0	1	0	0	1	0	**1**	0	1
C_5	1	0	0	0	1	0	0	1	0	1	**1**
C_6	0	1	1	0	0	0	**1**	0	0	1	1
C_7	0	1	0	1	0	0	0	1	1	**1**	0
C_8	**0**	1	0	0	1	1	0	0	1	0	1
C_9	0	0	1	1	**0**	1	0	1	0	0	1
C_{10}	0	0	1	0	1	0	1	**1**	1	0	0
C_{10}	0	**0**	0	1	1	1	1	0	0	1	0

Table 6.1: A 2-$(11, 5, 2)$ design with RTD $= 3$ and NCTD $= 1$.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
C_1	**1**	1	1	1	1	1	1	0	0	0	0	0	0	0	0
C_2	1	1	1	**0**	0	0	0	1	1	1	1	0	0	0	0
C_3	1	1	1	0	0	0	0	0	0	0	0	**1**	1	1	1
C_4	1	0	0	1	1	0	0	**1**	1	0	0	1	1	0	0
C_5	1	0	0	1	1	**0**	0	0	0	1	1	0	0	1	1
C_6	1	0	0	0	0	1	1	1	1	0	0	0	0	**1**	1
C_7	1	**0**	0	0	0	1	1	0	0	1	1	1	1	0	0
C_8	0	1	0	1	0	1	0	1	0	**1**	0	1	0	1	0
C_9	0	1	0	1	**0**	0	1	1	0	0	1	0	1	0	1
C_{10}	0	1	**0**	0	1	1	0	0	1	0	1	1	0	0	1
C_{11}	0	1	0	0	1	0	1	0	1	1	0	0	**1**	1	0
C_{12}	0	0	1	1	0	1	**0**	0	1	0	1	0	1	1	0
C_{13}	0	0	1	1	0	0	1	0	**1**	1	0	1	0	0	1
C_{14}	0	0	1	0	1	1	0	1	0	1	0	0	1	0	**1**
C_{15}	**0**	0	1	0	1	0	1	1	0	0	1	1	0	1	0

Table 6.2: A 2-$(15, 7, 3)$ design with RTD $= 3$ and NCTD $= 1$.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}
C_1	**1**	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
C_2	1	1	1	1	0	**0**	0	0	0	1	1	1	1	1	0	0	0	0	0
C_3	1	1	1	0	1	0	0	0	0	1	0	0	0	0	**1**	1	1	1	0
C_4	1	1	0	0	0	1	1	0	0	0	1	**1**	0	0	1	1	0	0	1
C_5	1	0	1	0	0	1	1	**0**	0	0	0	0	1	1	0	0	1	1	1
C_6	1	0	0	1	1	0	0	1	0	0	1	1	0	0	0	0	**1**	1	1
C_7	1	0	**0**	1	1	0	0	0	1	0	0	0	1	1	1	1	0	0	1
C_8	1	0	0	0	**0**	1	0	1	1	1	1	0	1	0	1	0	1	0	0
C_9	1	0	0	0	0	0	1	1	1	1	0	1	0	1	0	1	0	**1**	0
C_{10}	0	1	1	0	0	0	0	1	1	0	1	0	0	1	1	0	0	1	**1**
C_{11}	0	1	0	1	0	1	0	1	**0**	1	0	0	1	0	0	1	0	1	1
C_{12}	**0**	1	0	1	0	0	1	0	1	0	0	1	1	0	1	0	1	1	0
C_{13}	0	1	0	0	1	1	0	0	1	**1**	0	1	0	1	0	0	1	0	1
C_{14}	0	1	0	**0**	1	0	1	1	0	0	1	0	1	1	0	1	1	0	0
C_{15}	0	0	1	1	0	1	0	1	0	0	0	1	0	**1**	1	1	1	0	0
C_{16}	0	0	1	1	0	0	1	0	1	1	1	0	0	0	0	**1**	1	0	1
C_{17}	0	0	1	0	1	1	**0**	0	1	0	1	1	1	0	0	1	0	1	0
C_{18}	0	**0**	1	0	1	0	1	1	0	1	0	1	1	0	1	0	0	0	1
C_{19}	0	0	0	1	1	1	1	0	0	1	**1**	0	0	1	1	0	0	1	0

Table 6.3: A 2-$(19, 9, 4)$ design with RTD $= 3$ and NCTD $= 1$.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}	x_{23}
C_1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	**0**	0	0	0
C_2	1	1	1	1	1	0	0	0	0	**0**	0	1	1	1	1	1	1	0	0	0	0	0	0
C_3	1	1	1	0	0	1	1	0	0	0	0	1	1	0	0	0	**0**	1	1	1	1	0	0
C_4	1	1	**0**	1	0	1	0	1	0	0	0	0	0	1	1	0	0	1	1	0	0	1	1
C_5	1	1	0	0	0	0	0	0	1	1	1	1	0	1	0	1	0	1	**0**	1	0	1	0
C_6	1	0	1	1	0	0	1	0	1	0	0	0	0	1	0	**0**	1	0	0	1	1	1	1
C_7	1	0	1	0	0	0	0	1	**0**	1	1	0	1	0	1	0	1	1	0	0	1	1	0
C_8	1	0	0	1	0	**0**	0	1	1	1	0	1	1	0	0	1	0	0	1	0	1	0	1
C_9	1	0	0	0	1	1	1	0	0	1	0	0	1	0	1	1	0	**0**	0	1	0	1	1
C_{10}	1	0	0	0	1	1	0	1	0	0	1	0	0	1	**0**	1	1	0	1	1	1	0	0
C_{11}	1	0	0	**0**	1	0	1	0	1	0	1	1	0	0	1	0	1	1	1	0	0	0	1
C_{12}	0	1	1	0	1	0	0	1	1	0	0	**0**	0	0	1	1	0	1	0	1	1	0	1
C_{13}	0	1	1	0	**0**	1	0	0	1	0	1	0	1	0	0	1	1	0	1	0	0	1	1
C_{14}	0	1	0	1	1	0	1	**0**	0	1	0	0	0	0	0	1	1	1	1	0	1	1	0
C_{15}	0	1	0	1	0	1	0	0	0	1	1	1	0	0	1	0	1	0	0	1	1	**0**	1
C_{16}	**0**	1	0	0	1	0	1	1	0	0	1	1	1	1	0	0	0	0	0	0	1	1	1
C_{17}	0	1	0	0	0	0	1	1	1	1	0	0	1	1	1	0	1	0	1	1	0	0	**0**
C_{18}	0	0	1	1	1	0	0	0	0	1	1	0	1	1	0	0	0	1	1	1	**0**	0	1
C_{19}	0	0	1	1	0	0	1	1	0	0	1	1	**0**	0	1	1	0	0	1	1	0	1	0
C_{20}	0	0	1	0	1	1	**0**	0	1	1	0	1	0	1	1	0	0	0	1	0	1	1	0
C_{21}	0	0	1	0	0	1	1	1	0	1	**0**	1	0	1	0	1	1	1	0	0	0	0	1
C_{22}	0	0	0	1	1	1	0	1	1	0	0	1	1	**0**	0	0	1	1	0	1	0	1	0
C_{23}	0	**0**	0	1	0	1	1	0	1	0	1	0	1	1	1	1	0	1	0	0	1	0	0

Table 6.4: A 2-$(23, 11, 5)$ design with RTD $= 3$ and NCTD $= 1$.

6.3 Algorithm

Algorithm 1 The recursive teaching dimension of concept class $\mathcal{C}$

Input: Concept class $\mathcal{C}$ over the domain $\mathcal{X}$ of size m.
Output: The recursive teaching dimension of concept class $\mathcal{C}$.

```
function IS_TEACHING_SET(tpl)
    n ← number of concepts
    for i in range(n) do
        i_col_vals ← All sample sets of size t, 1 ≤ t ≤ m, for the row i^th
        for j in range(n) do
            j_col_vals ← All sample sets of size t, 1 ≤ t ≤ m, for the row j^th
            if  i_col_vals ≠ j_col_vals then i_col_vals is a teaching set.
            end if
        end for
    end for
    return The teaching sets corresponding to the i^th concept.
end function

function TDMIN(𝒞)
    for i in range(0,m + 1) do
        combs ←  All combinations of size i of set {1, . . . , m}
    end for
    for tpl in combs do
    Found the first teaching set associated to a concept by IS_TEACHING_SET(tpl)
    end for
    return The minimum teaching set corresponding to the i^th concept.
end function

function RTD(𝒞)
    n ← number of concepts
    while n > 0 do
        𝒞_m ← The concepts C with the minimum teaching dimension m
        Add m to a set called minset
        𝒞 ← 𝒞 \ 𝒞_m
    end while
    return The maximum element of minset.
end function
```